

DYSERTACJE
WYDZIAŁU
NEOFILOLOGII
UAM
W POZNANIU
(NOWA SERIA)

18

JĘZYKOZNAWSTWO

YURY FEDORUSHKOV

CONTEXTUAL
REPOSITORY
OF DICTIONARY
AGNONYMS IN
THE RUSSIAN
LOCAL PRESS
(1996–2006)

10,000 OF ENTRIES IN
THE HTML FORMAT

CONTEXTUAL REPOSITORY OF DICTIONARY
AGNONYMS IN THE RUSSIAN LOCAL PRESS
(1996–2006)

REPOZYTORIUM KONTEKSTOWE
AGNONIMÓW SŁOWNIKOWYCH
W ROSYJSKIEJ PRASIE LOKALNEJ
(1996–2006)

DYSERTACJE
WYDZIAŁU NEOFILOLOGII
UAM W POZNANIU
(NOWA SERIA)

18
JĘZYKOZNAWSTWO

YURY FEDORUSHKOV

CONTEXTUAL REPOSITORY OF DICTIONARY
AGNONYMS IN THE RUSSIAN LOCAL PRESS
(1996–2006)

10,000 OF ENTRIES IN THE HTML FORMAT

REPOZYTORIUM KONTEKSTOWE
AGNONIMÓW SŁOWNIKOWYCH
W ROSYJSKIEJ PRASIE LOKALNEJ
(1996–2006)

10,000 HASEŁ W FORMACIE HTML



POZNAŃ 2024

Projekt okładki:
Wydawnictwo Rys

Na okładce:
Fragment fasady Biblioteki
Wydziału Neofilologii UAM w Poznaniu

Recenzja wydawnicza:
dr hab. Henryk Jaroszewicz, prof. UW

Copyright by:
Yury Fedorushkov

Copyright by:
Wydawnictwo Rys

Korekta tekstu angielskiego
Graham Crawford

Korekta tekstu polskiego:
Sebastian Surendra

**Redakcja Naukowa
Dysertacji Wydziału Neofilologii UAM:**

Przewodnicząca:
prof. UAM dr hab. Dominika Skrzypek, prodziekan ds. nauki

Członkowie:
prof. UAM dr hab. Katarzyna Klessa
prof. UAM dr hab. Sylwia Adameczak-Krysztofowicz
prof. UAM dr hab. Barbara Łuczak
prof. zw. dr hab. Piotr Muchowski
prof. UAM dr hab. Wawrzyniec Popiel-Machnicki
prof. UAM dr hab. Krzysztof Stroński
prof. UAM dr hab. Janusz Taborek
prof. UAM dr hab. Władysław Zabrocki

Wydanie I, Poznań 2024

ISBN 978-83-68006-68-1

**DOI
10.48226/dwnuam.978-83-68006-68-1_2024.18**

Wydanie:



Wydawnictwo Rys
ul. Kolejowa 41
62-070 Dąbrówka
tel. 600 44 55 80
e-mail: tomasz.paluszynski@wydawnictworys.com
www.wydawnictworys.com

Table of contents

Foreword	9
I. Dictionary agnonyms.....	13
II. Scientific objective for the establishment of the repository	15
III. CRPLOM body and constraints for RDA.....	19
IV. Research methods.....	43
V. Verification of dictionary agnonyms.....	47
VI. Context and definition.....	49
VII. Excerpt criteria	51
VIII. Historical period: the decade for the "language of newspapers"	57
IX. Research opportunities in morphology.....	69
X. The discreteness of the word and the inadequacies of the RDA.....	73
XI. Structure of the RDA and access to the resource	79
XI.1. The vertical dimension of the RDA	81
XI.2. The horizontal dimension of the RDA.....	83
Bibliography	87
Literature	87
Verification dictionaries.....	95
Corpora and other textual resources.....	98
Index of tables.....	99
Index of illustrations	99

Spis treści

Przedmowa.....	103
I. Agnonim słownikowy.....	107
II. Cel naukowy utworzenia repozytorium	109
III. Korpus KRPLOM i ograniczenia dla RAS.....	113
IV. Metody badawcze.....	139
V. Weryfikacja agnonimów słownikowych.....	143
VI. Kontekst oraz definicja	145
VII. Kryteria ekscerpcyjne.....	147
VIII. Okres historyczny: dziesięciolecie millenium dla „języka gazet"	153
IX. Możliwości badawcze w zakresie morfologii.....	165
X. Dyskretność słowa oraz niedoskonałości RAS.....	169
XI. Struktura RAS oraz dostęp do zasobu	175
XI.1. Wymiar pionowy RAS	177
XI.2. Wymiar poziomy RAS	179
Bibliografia	183
Literatura naukowa.....	183
Słowniki weryfikacyjne.....	190
Korpusy i inne zasoby tekstowe.....	193
Spis tabel.....	195
Spis rysunków.....	195

CONTEXTUAL REPOSITORY OF DICTIONARY
AGNONYMS IN THE RUSSIAN LOCAL PRESS
(1996–2006)

10,000 OF ENTRIES IN THE HTML FORMAT

I would like to thank Mateusz Lango, PhD (Poznan University of Technology, Faculty of Informatics and Telecommunications, Institute of Informatics) for his help in solving programming problems, Maciej Waraczewski, PhD (director of the Poligłoci Language Centre in Poznań) for his help in contextualising the agnonyms, the CLARIN team for their assistance with the publication of the newspaper corpus – in particular to Tomasz Naskręt, MSc, and Jan Wierzchorek, PhD, IT specialists Maciej Danielak and Michał Rachwalski from the Faculty of Neophilology at Adam Mickiewicz University for their guidance with the PYTHON script code and their help with the graphic processing of the repository file in HTML format.

Foreword

A well-known fact pointed out by researchers is that one of the current lexical features of Slavic languages is their internationalisation (Mietła 2001: 52, Kaliszan 1980, Waszakowa 2005: 10, 26, 28, 206). In the literature on the subject, the openness of Russian, like that of other Slavic languages, is primarily discussed (Vashakova 2005: 194; cf. chapter title: *Tendency towards internationalisation as a Slavic-wide phenomenon*) to other languages (Trofimova 2001a, Alatortseva 1999). The researcher S.I. Alatortseva believes that neologisms are primarily a manifestation of internal cultural factors: "(...) The process of neologisation is inherent to a greater or lesser extent in all national languages, but each national culture has its own special attitude to it. Russia, in our opinion, is one of the countries where the attitude to linguistic innovation is too democratic. The main problem of this direction is national culture and foreign language neologisms." (Alatortseva 1999: 9). The lexis becomes richer and modified through the process of borrowing. New lexis appears, not previously recorded in texts. These new lexemes are precisely the litmus tests for highlighting the course of linguistic processes. There is, therefore, a clear need for a neographic lexical resource to enable the study of both the process of internationalisation of the Russian language in connection with the aforementioned 'opening' of Russia to the West during the political regime change of 1996–2006, and scientifically grounded research activities directed towards a strictly linguistic analysis. At this stage, this openness of the Russian Federation has undoubtedly been undermined by Russia's isolation from Western countries, the closure of its borders, the withdrawal of Western companies and corporations, the designation of Western countries as 'unfriendly countries', Russia's military interventions in Chechnya and later in Georgia, Syria, African countries and, on 24 February 2022, Ukraine. A large proportion of periodicals listed in our textual raw material – the corpus of Russian newspapers – have disappeared or adopted propaganda rhetoric. Today's Russia is dominated by media – including newspapers – subordinated to the language and rhetoric of power. From this point of view, it is worth

making available both a repository of previously unrecorded words with the contexts of their use, and the corpus of newspapers itself, where dialogue takes place and freedom of thought and speech, distinct from the monologue of the modern propaganda machine, is still noticeable. The public release of such a resource will make it possible to launch a whole series of research works related to the development of the lexis, morphology, vocabulary and semantics of modern Russian.

The repository of dictionary agnonyms (RDA) is is chronological with respect to both language and geography. It is aimed at a wide range of linguists – primarily lexicographers and lexicologists. The word sources are the archives of the local Russian press, from which *the Corpus of Russian Local Press of the Millennium Period* (CRLPMP) has been constructed. The study presents the result of the research and excerpting work exclusively within the corpus. It should be mentioned here that this corpus has been made public: CRLPMP has been made available in the DSpace system of the CLARIN-PL resource (Fedorushkov 2023).

I am submitting the author's selection of 10,000 keyword articles illustrating primarily the neologisms of the Russian local press from 1996–2006 collected in the various regions of the Russian Federation. The search of periodicals, including archives, was done on the basis of the 'printed version = online version' analogy, i.e. the printed issue of a particular newspaper with its corresponding online version. A corpus of texts was created from the online versions of the newspapers. The composition of the CRPLOM took place between 2005 and 2006. The year 2006 is therefore also the lower time limit for DA excerpting and any verification of the resulting words.

A preliminary list (LNS) of neologisms conceived in this way appeared as part of the algorithmised excerpting activities described in the doctoral dissertation entitled *Methods of excerpting Russian language neologisms (on the material of the contemporary Russian press)*¹ (Fe-

¹ The content of the dissertation can be accessed online at the Adam Mickiewicz University (UAM) Repository resource in Poznań – AMUR: (<https://repozytorium.amu.edu.pl> – WEB: 21.03.2021). The dissertation was written under the ac-

dorushkov 2008). Manual excerpting of new words is a very complicated process. This is related to a simple fact: the excerpting of available textual quantities is not feasible manually. A methodical apparatus is required to automate some of its steps. Therefore, a few technical terms will appear below mainly to denote the evolution of the LNS-list to the DA-list (cf. Chapter IV. *Research methods*). In order to extract neologisms, it was necessary to create a technology to get a list of graphically unique words – the so-called neonyms (graphic-lexical neologisms). The overall word list of the corpus is 250 million specimens. The final list of graphically unique words – the LNS(AK) list (cf. *ibid.*) – counts 300,000 strings of unique – non-lemmatised types. The selection of words in the repository's index can be described as original (authorial); agnonyms were selected and verified subjectively. The words comprising the LNS(AK) list were manually selected from the LNS-List. If the number of words with the prefix анти-* in the repository is a few dozen, then in one of the first DA lists words with анти-, -анти and -анти- can constitute a separate dictionary – there are almost 2,700 of them. However, this requires separate lexicographical operations. The process of DA excerpting can be expressed by the following sequence:

- CRLPMP (250 million tokens);
- LNS(AK) list (300,000 uniques absent from the Zaliznyak 1987 dictionary – cf. chapter entitled IV. *Research methods*);
- LNS-List (a list of more than 15,000 lemmatised DA without any context of occurrence, after verification of quotation in appropriately selected lexicographical sources). Examples of words from the LNS-List:

100-квартирный, 3D-видеоадаптер, АА-батарея, АВИА-ШОУ, АВИСМА-МетрАТек, Аз-Буки-Веди, АмурЭкспо-Форум-2006, о-о-чень, рагга-джангловый, ракетоподобный, рэпообразный, сайтотворчество, свинговать, тарантинофи-лософия, "телекамикадзе", трендмейкер, ультраалалайтовый, ультраэпатажный, файлообменный, фильмпроверщица etc.;

ademic supervision of Prof. Piotr Wierzchoń at the Institute of Linguistics at the Adam Mickiewicz University (now the Department of Ethnolinguistics). I would like to point out that this publication will make use of a small number of extracts of content from the dissertation in question – particularly in Chapter VIII.

- IDA index (an *index* of more than 10,000 DA verified for correctness in contexts of use and additionally contextualised – by the addition of a single sentence or longer passage from a context of use to a headword).

The content of the RDA is the result of contextualising the entire index of lemmas – more than 10,000 unique words highlighting the relevant lexeme in context, indicating the name of the periodical, the issue, the year of publication, and – most importantly: the region and city for each excerpt. I refer here to the opinion of an expert in the field of linguo-chronologisation: Piotr Wierzchoń, one of the co-authors of the ten-volume *Depository of the Lexicon of the Polish Language*, points out that the linguo-chronologisation excerption of neonyms leads to the next necessary parameters: the entry, the textual illustration, the dating, the geographical and temporal location (cf. Wierzchoń, ed. 2010).

Given the enormity of the technical, methodological, substantive nuances involved in the RDA, I have limited myself to including a detailed description of the creation of the RDA in the following chapters. I will describe first of all the specific purpose of the creation of the RDA, the general criteria for the excerpting of headwords, the stages of work on the RDA, the criteria for the selection of context and the process of contextualisation, the question of the definition of headwords, the corpus of newspapers, the historical period in which the periodical numbers were produced, the structure of the RDA, the structure of the headword article and the series of lexicographical decisions during the creation of the *a fronte* index, as well as a prolegomenon for the description of general trends in the development of the morphology of new words.

I. Dictionary agnonyms

In his book entitled *Photodocumentation. Chronologisation. Emen-dation. Theory and practice of verifying lexical material in linguistic research* (Wierzchoń 2008), Piotr Wierzchoń compiles a list of rules (so-called *directives*) for the chronologisation of vocabulary, pointing out that the term neologism should be used with great caution in order to avoid a false dating hypothesis. One of the directives states: "In order to minimise the number of false dating hypotheses, the use of the linguo-chronological term neologism should be abandoned in relation to language units considered to be dated to the post-war period in favour of the term dictionary agnonym, which states that the language unit in question was not recorded in post-war dictionaries or lexicographical studies" (Wierzchoń 2008: 120). In view of the given objection, instead of the terms *neologism* or "more cautious" – *potential neologism* – or even *neonym*, I will use the superior term, namely *dictionary agnonym* (DA). By using this hyperonym, I am emphasising that a given unit of language has not been recorded in a particular series of dictionaries preceding the creation of the list, in which more than 10,000 DA have been accumulated. In other words, it is a technical and strategic term. Within its functional scope are not only neonyms (neologisms, occasionalisms), but also classically understood agnonyms (i.e. rare, little-known, little-used, unintelligible – extremely peripheral words), historicisms and hapax legomena – occasional authorial words used by individual journalists, among others, to intensify expression. Of the DA, it is the new words, generated or repeated for contemporary purposes, that predominate to a greater extent, such as: *SMS-ругательство, смс, онлайн-болтовня, фэшн-неделя, FM-станция, аудиовидеосинхронизация, профайлер, путиноманья, путиномика, брендмейкер, порноюмор*. A neologism can be registered just among a number of dictionary agnonyms. A neologism is a usual unit: it can be considered new if it lexicalises and actively functions in the usus, its non-peripheral areas only after the 'test of time' according to the idea contained in the definition of neologism by Russian neology researcher Nadezhda Kotyolova. These are 'words, word meanings, idioms that

usually exist in a certain period, in a certain language, sub-language, linguistic sphere and do not exist in a certain limited previous period in the same language, sub-language, linguistic sphere' (Kotyolova 1978: 22). A neologism, in view of this, is a word that 'intends' to cease to function as a new word, becoming a word without the connotative supraseme 'new', a serial, active word. So far, however, it contains such a seme, due to which it is considered new lexis. To reiterate: if we want to recognise a word as a neologism, we should examine its fate in the contexts of use, its frequency, territorial, temporal, dialectical, socio-lectic range. The recognition of such a word as a neologism begins with the recognition of such a word as a dictionary agnonym. It is this idea that guides the formulation of the purpose of the repository.

II. Scientific objective for the establishment of the repository

The lexis in RDA dates from the period 1996–2006, which can technically be described as the decade. There is a view (Misturska-Bojanowska 2013, Bortnowska 2016 et al.) that the inflectional language of Russian in the period of development under study (late 20th – early 21st century Russia – cf. Bieleń 1997, Dravich 1997, Mandrikova, Krotova 2013) intensively continues the drive towards analyticisation and agglutination, with this tendency concerning vocabulary from a specific thematic spectrum (Handke 1976, Maltchina 1974, Pospiszyl 2000). This is especially true for composites (Yodlovsky 1936, Bukchina, Kalakutkaya 1974, Vashakova 2015, Vashakova 2021, especially for Russian Gabdreeva, Gurchiani 2012, Sinyachkina 1997, Vepreva 2005). The trends of analyticisation and paradigmatisation in contemporary Russian morphotactics represent a stage of enrichment of specific lexical spheres and contribute to a change in the direction of language development. The enrichment of the language with new vocabulary – neonyms – is precisely the area that signals these changes. In order to provide insight into the indicated trend of the growth of neonyms with specific parameters, in 2009 I started composing and contextualising the IDA – an index of dictionary agnonyms and their graphic variations (cf. Bańko, Hebal-Jeziarska 2015) in the process of adaptation. In 2022, I completed the dating verification activity and carried out context selection for each word in the IDA index. In addition to the context, each DA was provided with metadata in addition to the context: title of the local newspaper, date (year), name of the region of occurrence of the DA and city with a geographical map. Parameters were collected for each of the ten thousand words in the IDA. For example, for the keyword *земфироваться* these are as follows:

Table 1. Example of key data in a keyword article in RDA

DATA TYPE	ELEMENTS OF THE KEYWORD ARTICLE
1. Lemmatised DA as a keyword highlighted in bold. If in context DA appears with inverted commas, the inverted commas remain in the slogan.	"земфироваться"
2. Context of use with the SA context form highlighted	Ю.Т. Положение в русском роке такое же, как и в роке вообще, то есть налицо уход в некие "потусторонние миры", не сказать, чтобы в подполье. Есть волна коммерческого рокака, которого все больше, все более мейнстримизуется (мейнстримизирует), "земфируется" , "муммитроляется", "ошнурируется" и так далее.
3. Metadata: title of journal, number and year of issue, Russian Federation sub-region (RFS), city	<Завтра №52 (632); 2005; Москва [as RFS]; Москва [as city]>.

Observing the word-neonyms from a given set, one can observe an exponential tendency towards (...) affixation (prefixation / prefixoidation / suffixation / suffixoidation), an increase in conjunctive constructions and compositions (e.g. two-core compounds with at least one borrowed element, e.g. *IT-технология*, *sms-голосование*, *лонг-дринк*, *интернет-переписка*, *смарт-фон*, *смарттелефон*). In view of this, these words can be classified – in possible studies – from the position of the most subtle classification – the so-called ‘pentagonal’ scale: affix – affixoid – semi-affixoid – radixoid – radix (radical, root) (Minina, Bartkov 1980, Bartkov 2006, Bartkov, Minina 2019) with a refined algorithm for distinguishing individual formants in functional terms.

Researchers indicate that the language of newspapers clarified in the 1970s–80s (Novikova 2002, Kostomarov 1971, Gogova 1975, Panov 1988). In the 1970s–1980s (Novikova 2002, Kostomarov 1971, Gogova 1975, Panov 1988), while in the millennium period it experienced its apogee – it even clarified a specific – liberalised, in opposition to the literary norm (Kostomarov 1999, Chubinsky 1997, Donaj 2001) – sociolect of local newspaper journalists, which survived the transfer from printed newspapers to digital versions, thus contributing to increased insight by citizens (Trofimova 2001b). The aim of the RDA repository is primarily to expose both the vocabulary itself and to look at its evolution. Furthermore, I find that during the period described, one can observe a phenomenally high growth, even an explosion of presumably recent words not observed in previous² dictionaries – cf. the infographic below:

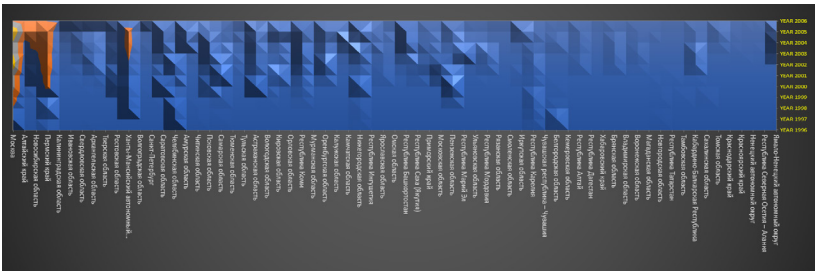


Figure 1. Explosion in the growth of unique words in the millennium decade (individual regions and years) [own study]

Subsequent research objectives posed to the linguistic data obtained should therefore take into account the intrinsic features of these data. I refer, of course, to the morphotactic properties of the words recorded in the RDA. Another important feature of these words is their neonymic uniqueness. This feature requires in-depth search and consultation verification in individual textual lexicons.

² By this I mean words not observed in the contents of the verification dictionaries. This list can be consulted in the subsection entitled *Verification dictionaries* in the Bibliography).

Let me make it clear once again: the aim of the work on RDA was not to determine, among other things, the vocabulary trends of modern Russian, nor did I focus only on the manifestations of internationalisation in the Russian vocabulary of the period under study. Of course, research 'updating' these issues is possible further down the line. However, these studies are conditioned by sufficiently extensive research material. The aim of the work on the RDA of the paper was to provide as abundant material as possible, not to analyse it. Certain elements of the analysis of modern Russian are naturally covered in Fedorushkov's 2008 dissertation, in which a description of a series of analytical decisions also took place. These decisions concerned the question of which atypical constructions can – with certain reservations, e.g. with the parameter *in status nascendi* – be regarded as acceptable in Russian, and which creations remain definitely incorrect; thus, for example, it seems that constructions of the type: *девушка*-"*молодогравдейка*", *интернет*-"*ливайс*", *культуртрегер*-"*Фаланстер*", *курсант*-"*вкиушник*", *куртка*-"*бомбер*", *муж*-"*анекдотоман*" etc.) should be considered correct. In conclusion, it should be firmly stated that the RDA in the above sense has a registration value, although in the practice of registration it is impossible to evade decisions of a normative nature.

What do the research methods and all the activities associated with them represent? Before answering this question, I will first describe the corpus of newspapers from which the verbal material was extracted. Let us then move on to discuss the statistical and geopolitical properties of the CRPLOM.

III. CRPLOM body and constraints for RDA

At the time of its creation (2005–2006³), the Russian Federation included 21 republics, 48 regions, 7 countries, 1 autonomous region, 7 autonomous districts, 2 cities of regional significance (St. Petersburg and Moscow): a total of 86 so-called *sub-regions* (RFSs) of federal significance⁴.

The resulting full CRLPMP word list for 2007 counted more than 5 million non-repeating word-types (*unique words, types*) and 250 million specimens (tokens). The corpus included materials and archives from the open public access of about 300 newspapers from 76 regions of the Russian Federation. No http equivalents of local newspapers could be found in the rest of the regions. The CRLPMP material consists of newspaper data from 1995 to 2006/early 2007. The archive data falling into the indicated years was not symmetrical: the archive of one newspaper – is data for one year, and of another – even for the entire millennium decade. The CRPLOM counts more than half a million asymmetrical http files: a single file can contain an entire issue of a newspaper, as well as a single news article or just the issue’s table of contents.

The local press of the CRLPMP consists of dailies, regional weeklies (the so-called областные, e.g. *Звезда Алтая, Карельская Губерния*), city newspapers (e.g. Архангельск, *Каргополье*) and central – capital city ones (St. Petersburg, Moscow: *Аргументы и Факты, Петербургский Час Пик, Российская газета*). These newspapers

³ Currently, the Republic of Crimea and the city of Sevastopol of regional importance have been annexed to Russia, with states, organisations, including the UN, not recognising the annexation of these territories to the integrated territories of the Russian Federation. The same applies to the Donetsk and Luhansk Republics.

⁴ Currently, the Russian Federation consists of 89 RFSs: 48 regions, 24 republics, 9 states, 3 cities of federal importance, 4 autonomous districts and 1 autonomous oblast. The affiliation of some RFSs to the Russian Federation is disputed: they are only recognised within the Russian Federation. This is related to state policy, e.g. Russia’s military invasion of Ukraine, the information-propaganda war in Donbas and Luhansk, and the annexation of Crimea. We will suggest the administrative division from the time of DA extraction for this repository until 2014.

illuminate the events of individual regions, cities, villages, i.e. purely local information on socio-political topics (e.g. politics, culture, sports, education, religion, ecology, crime news, etc.).

For the 2006 I found 2,302 central and regional periodicals, having both print and online versions. From the Internet it was possible to download only 280 newspapers having online equivalents, a readable html structure of the pages, i.e. ‘viable’ due to the feasibility of purely technical processing (e.g. change of font standard coding, possibility of making copies, etc.). The essential parameter was the active server of the specific newspaper at the time of data collection (2006, early 2007). Prior to 2006, 86 FR subs were recognised. The 280 newspapers mentioned were from 71 sub-areas – an LNS, i.e. an initial list of new words, was created on the textual material of these newspapers. In this Repository, the dictionary agnonyms come from 235 periodicals, which in turn come from 68 Russian sub-regions. This reduction in the number of Russian regions has to do with the real result of the excerpting and the modification of the LNS.

The hard drive on which the CRLPMP is housed has more than 650,000 text files. These range from individual newspaper articles to entire pages (in Russian terminology газетная полоса), containing anything from a single article to the content of an entire issue (cf. Table 2).

In the selection of newspapers for the CRPLOM, most attention was focused on the local press. This press focuses on events ‘closer’ to the residents of a particular region, i.e. purely local news with socio-political themes (e.g. culture, sport, education, religion, ecology, letters to the editor, crime news, etc.). Particular attention was paid to dailies, publishing articles in the style of small news⁵ newspapers. In Table 2 I present a representative inventory of the periodicals used in this repository. I consider this census – openly and in its entirety provided here – to be one of the most significant assets of this study⁶. The limitations of the corpus boil down to the figures: of the 280 newspapers of the CRLPMP, 235 periodicals from 93 cities located in

⁵ Dailies are considered (e.g. Kurcz et al. 1990) to be the best source of small press news.

⁶ In the practice of linguistic research, it is all too often the case that general, not to say vague, information is given, completely unverifiable, concerning only a selection of press titles subjected to random excerpting.

68 RFSs were used in the Repository – cf. Table 2. The following data are included in the table:

Column 1 One type of data: alphabetical list RFS.

Column 2. Three types of data:

- name of the RFS in 2007;
- total number of journals used in IDA;
- total number of DA used in the LNS-List.

Column 3 One type of data: list of RFSs in 2022 indicating changes in number and names of RFSs.

Column 4 One type of data: name of RFS capital in 2022.

Column 5 Two types of data:

- RFS city (in 2007);
- number of newspapers.

Column 6. 2 types of data:

- magazine name;
- number of DA on the LNS-List for any journal.

Table 2. Linguoepochronological data in RDA: RFS and DA

1	2	3	4	5	6
	Name of RFS in 2007; total number of journals used in IDA; total number of DA used	List of RFSs in 2022, indicating changes in RFS numbers and names	Name of RFS's capital in 2022	City of RFS (for 2007); number of journals used in RDA	Name of periodical; number of DA extracts in IDA for each periodical
1.	Алтайский край; 6; 473	Алтайский край	Барнаул	Барнаул; 6	Altay Daily Review; 131 Ваше дело; 30 Комсомольская правда – Алтай; 21 Молодёжь Алтай; 20 Монитор+ (magazine); 240 Свободный курс; 31
2.	Амурская область; 3; 161	Амурская область	Благовещенск	Благовещенск; 2	Амурская правда; 145 Комсомольская правда – Приамурье; 15 ТВС-анонс; 1

1	2	3	4	5	6
3.	Архангельская область; 12, 140	Архангельская область	Архангельск	Архангельск; 6	Архангельск; 11 Архангельская лесная газета; 4 Ведомости Поморья; 1 Волна; 15 Дачная; 2 Двинская правда; 5
				Шенкурск; 1	Важский край; 6
				Каргополь; 1	Каргополье; 3
				Северодвинск; 4	Завалинка; 3 Корабельная сторона; 58 Северный рабочий; 27 Травинка; 5
4.	Астраханская область; 3; 227	Астраханская область	Астрахань	Астрахань; 3	Комсомолец Каспия; 2 Волга; 67 Астраханские ведомости; 158
5.	Белгородская область; 3; 37	Белгородская область	Белгород	Белгород; 2	Житьё-Бытьё; 15 Миссионерское обозрение (magazine); 1
				Старый Оскол; 1	Старый Оскол; 21

1	2	3	4	5	6
6.	Брянская область; 2; 27	Брянская область	Брянск	Брянск; 2	Аргументы и Факты – Брянск; 5 Брянское время; 22
7.	Владимирская область; 1; 2	Владимирская область	Владимир	Владимир; 1	Владимирские ведомости; 2
8.	Волгоградская область; 3; 47	Волгоградская область	Волгоград	Волгоград; 3	Волга и Дон; 1 Волгоградская правда; 34 Комсомольская правда – Волгоград; 12
9.	Вологодская область; 4; 126	Вологодская область	Вологда	Вологда; 4	Вологодская неделя; 1 Комсомольская правда – Вологда; 15 Красный Север; 3 Премьер; 107
10.	Воронежская область; 2; 17	Воронежская область	Воронеж	Воронеж; 2	Берег; 14 Подъем (magazine); 3
11.	Еврейская автономная область; 0; 0	Еврейская автономная область	Биробиджан	N/A	N/A
12.	Ивановская область; 2; 120	Ивановская область	Иваново	Иваново; 2	Частник°; 119 Хронометр – Иваново; 1

1	2	3	4	5	6
13.	Иркутская область; 1; 36	Иркутская область	Иркутск	Иркутск; 2	Аргументы и факты – Восточная Сибирь; 36
14.	Кабардино-Балкарская Республика; 1; 48	Кабардино-Балкарская Республика	Нальчик	Нальчик; 1	Газета Юга; 48
15.	Карачаево-Черкесская Республика; 0; 0	Карачаево-Черкесская Республика	Черкесск	N/A	N/A
16.	Калининградская область; 2; 83	Калининградская область	Калининград	Калининград; 2	Калининградская правда; 82 Янтарный Караван; 1
17.	Калужская область; 1; 54	Калужская область	Калуга	Калуга; 1	Весть; 54
18.	Камчатская область (until 2006); 1; 63	Камчатский край	Петропавловск-Камчатский	Петропавловск-Камчатский; 1	Камчатское время; 63
19.	Кемеровская область (until 2019); 2; 4	Кемеровская область – Кузбасс	Кемерово	Кемерово; 1 Новокузнецк; 1	Томь; 1 Кузнецкий тракт; 3
20.	Кировская область; 3; 134	Кировская область	Киров	Киров; 3	Комсомольское племя; 1 Вятский край; 128 Вятский наблюдатель; 5

1	2	3	4	5	6
21.	Костромская область; 0; 0	Костромская область	Кострома	N/A	N/A
22.	Краснодарский край; 1; 1	Краснодарский край	Краснодар	Сочи; 1	Курортный проспект; 1
23.	Красноярский край; 1; 1	Красноярский край	Красноярск	Норильск	Заполярная правда; 1
24.	Курганская область; 0; 0	Курганская область	Курган	N/A	N/A
25.	Курская область; 0; 0	Курская область	Курск	N/A	N/A
26.	Ленинградская область; 0; 0	Ленинградская область	Санкт-Петербург, Гатчина	N/A	N/A
27.	Липецкая область; 0; 0	Липецкая область	Липецк	N/A	N/A
28.	Магаданская область; 1; 5	Магаданская область	Магадан	Магадан; 1	Вечерний Магадан; 5

1	2	3	4	5	6
29.	Москва; 28; 2371	Москва	Москва	Москва; 28	<p>Авторевю; 1 Вестник МСПС (magazine); 1 Гудок; 33 Завтра; 1373 Молоко (magazine); 11 Московская переспектива; 94 Комсомольская правда – Москва; 1 Интерфакс Время; 16 Московские новости; 560 Наш современник (magazine); 1 Независимая газета; 4 Новая газета; 2 Новые Известия; 3 Огонёкъ (magazine); 1 Подвиг (magazine); 2 Полъём (magazine); 28 Полдень (magazine); 1 Россия; 31 Российская Бизнес-газета; 1 Российская газета; 3 Русская Газета; 1</p>

1	2	3	4	5	6
					<p>Русская жизнь (magazine); 3 Русский курьер; 63 Сети/Network world (magazine); 1 Слово (magazine); 1 Совершенно секретно; 102 Современник; 1 Хронос (regidic stock); 29 Элитный персонал; 1</p>
30.	Московская область; 7; 100	Московская область	Москва, Красногорск	<p>Ногинск; 2 Щёлково; 1 Дмитров; 1 Сергиев Посад; 1 Серпухов; 2</p>	<p>Богородские вест; 11 Богородский край; 1 Время; 14 Дмитровский Вестник; 19 Новое Зеркало – Сергиев Посад; 1 Серпуховские вест; 43 Серпуховские новости; 11</p>
31.	Мурманская область; 27; 26	Мурманская область	Мурманск	Мурманск; 1	Мурманский Вестник; 27
32.	Ненецкий автономный округ; 1; 2	Ненецкий автономный округ	Нарьян-Мар	Нарьян-Мар; 1	Известия НАО; 2

1	2	3	4	5	6
33.	Нижегородская область; 3; 155	Нижегородская область	Нижний Новгород	Нижний Новгород; 2 Саров; 1	Курс Н; 52 Нижегородские новости; 1 Городской Курьер; 102
34.	Новгородская область; 3; 5	Новгородская область	Великий Новгород	Великий Новгород; 3	Волхов; 2 Новгородские ведомости; 1 Торговая сторона; 2
35.	Новосибирская область; 5; 622	Новосибирская область	Новосибирск	Новосибирск; 5	Честное слово; 36 Сибирские огни (magazine); 36 Континент Сибирь; 547 Вечерний Новосибирск; 2 Комсомольская правда – Новосибирск; 1
36.	Омская область; 1; 12	Омская область	Омск	Омск; 1	Комерческие вести; 12
37.	Оренбургская область; 3; 143	Оренбургская область	Оренбург	Оренбург; 2	Яикь; 1 Вечерний Оренбург; 4
38.	Орловская область; 2; 71	Орловская область	Орёл	Орск; 1 Орёл; 2	Орская Хроника; 138 Просторы России; 69 Орловский Меридиан; 2

1	2	3	4	5	6
39.	Пензенская область; 3; 157	Пензенская область	Пенза	Пенза; 3	Аргументы и факты – Пенза; 30 Молодой ленинец; 115 Московский комсомолец – Пенза; 12
40.	Пермская область (Пермский край – since 2005); 10; 1078	Пермский край	Пермь	Пермь; 10	Enter-Профи (magazine); 18 Звезда; Звезда-online; 274 Капитал-Weekky; 65 Комсомольская правда – Пермь; 14 Местное время – Пермь; 44 Новости ПГУ; 3 Новый компаньон; 366 Пермские новости; 263 Пермский университет; 27 Хронометр – Пермь; 4
41.	Приморский край; 1; 74	Приморский край	Владивосток	Владивосток; 1	Арсеньевские вести; 74
42.	Псковская область; 2; 229	Псковская область	Псков	Псков; 2	Псковская губерния; 219 Псковская Правда; 10
43.	Республика Адыгея; 0; 0	Республика Адыгея	Майкоп	N/A	N/A

1	2	3	4	5	6
44.	Республика Алтай; 2; 26	Республика Алтай	Горно-Алтайск	Горно-Алтайск	Постскриптум; 24 Правда; 2
45.	Республика Башкортостан; 6; 152	Республика Башкортостан	Уфа	Уфа; 6	Бельские просторы (magazine); 106 Уфимский меридиан; 1 Уфимские новости; 1 Йэшлек; Молодость; 36 Версия Башкортостана; 4 Московский комсомолец – Уфа; 4
46.	Республика Бурятия; 0; 0.	Республика Бурятия	Улан-Удэ	N/A	N/A
47.	Республика Дагестан; 1; 3.	Республика Дагестан	Махачкала	Махачкала; 1	Новое дело; 3
48.	Республика Ингушетия; 1; 52	Республика Ингушетия	Магас	Назрань	Сердало; 52
49.	Республика Калмыкия; 0; 0	Республика Калмыкия	Элиста	N/A	N/A
50.	Республика Карелия; 6; 76	Республика Карелия	Петрозаводск	Петрозаводск; 6	Площадь Свободы; 1 Петрозаводск; 4 ТВР-Панорама; 3 Моя газета; 4 Лицей; 9 Карельская Губерния; 55

1	2	3	4	5	6
51.	Республика Коми; 5; 301	Республика Коми	Сыктывкар	Сыктывкар; 4	Молодежь Севера; 141 Зырянская жизнь; 66 Комсомольская правда – Сыктывкар; 39 Verbum; 7 НЭП+С; 48
52.	Республика Марий Эл; 2; 67	Республика Марий Эл	Йошкар-Ола	Йошкар-Ола; 2	Марийская правда; 1 Московский комсомолец – Марий Эл; 66
53.	Республика Мордовия; 2; 30	Республика Мордовия	Саранск	Саранск; 1 Ардатов; 1	Вечерний Саранск; 7 Маяк; 23
54.	Республика Саха (Якутия); 2; 56	Республика Саха (Якутия)	Якутск	Якутск; 2	Якутия; 13 Илкэн; 34
55.	Республика Северная Осетия – Алания; 1; 2.	Республика Северная Осетия – Алания	Владикавказ	Нерюнгри; 1 Владикавказ; 1	Индустрия Севера; 9 Дарьал (magazine); 2
56.	Республика Татарстан; 3; 3	Республика Татарстан	Казань	Казань; 2 Набережные Челные; 1	Вечерняя Казань; 1 Республика Татарстан; 1 Челны ЛТД; 1

1	2	3	4	5	6
57.	Республика Тыва; 0; 0	Республика Тыва	Кызыл	N/A	N/A
58.	Республика Хакасия; 0; 0	Республика Хакасия	Абакан	N/A	N/A
59.	Ростовская область; 5; 208	Ростовская область	Ростов-на-Дону	Ростов-на-Дону; 4	Приазовский край/ Стадион; 1 Комсомольская правда – Ростов-на-Дону; 51 Город N; 142 Газета Дона; 1
60.	Рязанская область; 1; 7	Рязанская область	Рязань	Волгодонск; 1 Рязань; 1	Вечерний Волгодонск; 13 Рязанские ведомости; 7
61.	Самарская область; 7; 185	Самарская область	Самара	Самара; 4	Самарская газета; 1 Ваш выбор; 13 Благовест; 49 Аргументы и факты – Самара; 48
				Сызрань; 1 Тольятти; 2	Гильдия; 2 Тольяттинское обозрение; 59 Площадь Свободы; 13

1	2	3	4	5	6
62.	Санкт-Петербург; 3; 55	Санкт-Петербург	Санкт-Петербург	Санкт-Петербург; 3	Петербургский Час-Пик; 10 Октябрьская магистраль; 44 Невское время; 1
63.	Саратовская область; 5; 270	Саратовская область	Саратов	Саратов; 5	Саратовский Арбат; 2 Саратовские вести; 93 Неделя области; 79 Богородские вести; 1 Богатей; 95
64.	Сахалинская область; 2; 3	Сахалинская область	Южно-Сахалинск	Южно-Сахалинск	Наши острова; 1 Сахалинское время; 2
65.	Свердловская область; 6; 143	Свердловская область	Екатеринбург	Екатеринбург; 6	Деловой квартал; 64 Деловой круг; 29 Депутатский круг; 5 Известия Уральского государственного университета (magazine); 1 Пенсионер; 17 Тайна жизни; 26
66.	Смоленская область; 2; 7	Смоленская область	Смоленск	Смоленск; 2	Рабочий путь; 6 Никольское кольцо – Смоленск; 1

1	2	3	4	5	6
67.	Ставропольский край; 0; 0	Ставропольский край	Ставрополь	N/A	N/A
68.	Тамбовская область; 3; 3	Тамбовская область	Тамбов	Тамбов; 2	Тамбовский курьер; 1 Тамбовский меридиан; 1
69.	Тверская область; 2; 60	Тверская область	Тверь	Мичуринск; 1 Тверь; 2	Мичуринская правда; 1 Тверские ведомости; 5 Вече Твери; 55
70.	Томская область; 0; 0	Томская область	Томск	Томск	Красное знамя; 15
71.	Тульская область; 2; 44	Тульская область	Тула	Тула; 2	Молодой коммунар; 7 Комсомольская правда – Тула; 37
72.	Тюменская область; 2; 58	Тюменская область	Тюмень	Тюмень; 2	Тюменские известия; 34 Тюменские ведомости; 24
73.	Удмуртская Республика; 0; 0	Удмуртская Республика	Ижевск	N/A	N/A
74.	Ульяновская область; 6; 30	Ульяновская область	Ульяновск	Ульяновск; 6	Деловое обозрение; 9 Аргументы и факты – Ульяновск; 3 Ульяновск сегодня; 1 Ульяновский меридиан; 3 Аргументы и факты – Ульяновск; 3 Вестник – Ульяновск; 13

1	2	3	4	5	6
75.	Хабаровский край; 1; 3	Хабаровский край	Хабаровск	Хабаровск; 1	Амурский Меридиан; 3
76.	Ханты-Мансийский автономный округ; Ханты-Мансийский автономный округ – Югра (from 2003); 9; 786	Ханты-Мансийский автономный округ – Югра	Ханты-Мансийск	Сургут; 6	Вестник – Сургут; 299 Вестник; приложение Синица – Сургут; 1 Контакт; 19 Московский комсомолец – Югра; 65 Сибирский газетчик; 9 Сургутская трибуна; 297
				Когалым; 1	Когалымские вести; 3
				Нижневартовск; 1	Местное время; 92
				Урай; 1	Знамя; 1

1	2	3	4	5	6
77.	Челябинская область; 11; 297	Челябинская область	Челябинск	Челябинск; 9	Южноуральская панорама; 45 Челябинский рабочий; 35 Лидер; 1 Злагоустовский рабочий; 13 Деловой Урал; 18 Городской Дилижанс; 37 Вечерний Челябинск; 20 Бизнес-класс; 3 Аргументы и Факты – Челябинск; 5
				Миасс; 1	Миасский рабочий; 118
				Магнитогорск; 1	Магнитогорский рабочий; 2
78.	Чеченская Республика (since 2003); 0; 0	Чеченская Республика	Грозный	N/A	N/A
79.	Читинская область; 3	Забайкальский край	Чита	Чита; 3	Экстра реклама; 49 Забайкальский рабочий; 27 Забайкальская новь; 8
80.	Чувашская республика – Чувашия; 1; 13	Чувашская Республика – Чувашия	Чебоксары	Чебоксары; 1	Советская Чувашия; 13

1	2	3	4	5	6
81.	Ямало-Ненецкий автономный округ; 1; 100	Ямало-Ненецкий автономный округ	Салехард	Ноябрьск; 1	Слово нефтяника; 100
82..	Ярославская область; 4; 37	Ярославская область	Ярославль	Ярославль; 3	Екклесиаст; 19 Золотое кольцо; 4 Юность; 10
83.	N/A; 0; 0	Севастополь	Севастополь	Рыбинск; 1	Рыбинская среда; 4
84.	Чукотский автономный округ	Чукотский Автономный округ	Анадырь	N/A	N/A
85.	N/A; 0; 0	Республика Крым	Симферополь	N/A	N/A
86.	Таймырский (Долгано-Ненецкий) автономный округ (until 2005); 0; 0	N/A (now: Таймырский Долгано-Ненецкий район Красноярского края)	Таймыр	N/A	N/A
87.	Агинский Бурятский автономный округ (until 2004); 0; 0	N/A (now it is Агинский Бурятский округ в составе Забайкальского края)	Агинское	N/A	N/A

1	2	3	4	5	6
88.	Коми-Пермяцкий автономный округ (until 2004); 0; 0	N/A (now it is: Коми-Пермяцкий округ Пермского края).	Кудымкар	N/A	N/A
89.	Эвенкийский автономный округ (until 2007); 0; 0	N/A (now: Эвенкийский район Красноярского края).	Тура	N/A	N/A
90.	Усть-Ордынский Бурятский автономный округ (until 2006); 0; 0	N/A (at present it is: Усть-Ордынский Бурятский округ Иркутской области)	Усть-Ордынский	N/A	N/A
91.	Корякский автономный округ (until 2006); 0; 0	N/A (now this is: Корякский округ Камчатского края).	Палана	N/A	N/A

1	2	3	4	5	6
92.	N/A; 0; 0	<p>* Республика Крым (as of 2014)</p> <p>* an administrative entity not recognised by most of the international community UN</p>	Симферополь	N/A	N/A
93.	N/A; 0; 0	<p>** Донецкая Народная Республика (from 2022).</p> <p>** an administrative entity not recognised by the majority of the UN international community</p>	Донецк	N/A	N/A

1	2	3	4	5	6
94.	N/A; 0; 0	<p>*** Луганская Народная Республика (since 2022)</p> <p>*** an administrative entity not recognised by the majority of the UN international community</p>	Луганск	N/A	N/A

The layout of the data in the table allows independent statistical calculations. The geographical coverage of IDA words is virtually the entire map of the Russian Federation:



Figure 2. Locations of IDA word occurrences on the map of regions [own study]

I do not provide internet addresses because of the huge series of purely historical modifications: since 2007, many newspapers have simply ceased to exist, periodicals have changed their addresses or even their names. Detailed information on address data and the number of resources can be found in Fedorushkov's 2008 dissertation (chapter entitled *Regional press*, 80–95).

IV. Research methods

Although the purpose of this Introduction to the repository is not to describe research methods and, in general, it is not a scientific monograph on methods and techniques for obtaining neologisms from a corpus of texts in an automated manner, I would like to sketch out the basic issues to eliminate some of the doubts about the suspected unrepresentativeness of the linguistic data.

In the most general terms, the range of methods and techniques in the study can be divided into two main stages: the collection of DA and the composition of RDA. Both the first and second stages are inextricably linked to parallel insights into corpus data. Therefore, the following will not be discussed in detail:

- corpus queries;
- the tools used to analyse the data and build the corpus and important details of the technical processes involved;
- software, algorithms, analytical methods that have been used to extract, process and organise DA;
- motives for the choice of words;
- lexicographical queries and verification of word presence;
- motives for word elimination: data cleaning, normalisation and handling dialectical varieties,
- ensuring the accuracy and usability of the repository;
- the methodological challenges and limitations of collecting and verifying the huge amount of data from regional newspapers;
 - The representativeness of the data and the usefulness of the RDA data for research into morphological trends.

All these issues are discussed in detail in my dissertation. I limit myself only to the highlights of the research method. In the following chapters of the Introduction, one can additionally learn about the criteria for lexicographic verification and the motives for selecting DA for RDA.

The research method of interest can be described as an algorithmised series of excerpting and verification steps. The path to arrive at the entire DA list can be illustrated by the excerpting and verification selection steps mapped in Table 3, which modifies the corpus word list

(LSK) to the DA index contained in this RDA. Additional commentary on the individual stages (obtained lists optimised to the final DA list) is provided below Table 3. The obtaining of specific lists is synonymous with the completion of the individual major stages.

Table 3. Path to the DA list

LIST OF ABBREVIATIONS	DESCRIPTION OF THE LIST
LSK	corpus word list
LSKA	list of corpus words with morphoanalysis data
LNS(AK)	„list of new words" + grammatical codes (grammemes and ancodes)
LNS	„list of new words" without homographs
LNS(K)	contextual word list
LNS(L)	list of lemmatised words
DA(BVD)	list of DA before verification in Russian dictionaries
DA(AVD)	DA list after verification in Russian dictionaries.
DA(BVC)	DA list before verification in context and supporting data
DA(AVC) = IDA	DA list after verification in context and supporting data

LSK. Obtaining all the words of the CRLPMP corpus. The tool to obtain – is the dtSearch Desktop indexer and segmenter (dtsearch.com) which allows character sequences to be split into tokens from space to space.

LSKA. Subjecting the words of the corpus to morphological analysis and obtaining the results of such analysis with the most important information: lemma of the word, code for the grammatical category cluster, information on the presence of words in the Zaliznyak 1987 dictionary. 1st stage of corpus queries. Using AOT technology for morphological analysis of Russian (aot.ru).

LNS(AK). Obtain a raw list of potential DA with grammatical codes to excerpt the exact wordforms of all parts of speech. The list includes homographs. The motivation for the selection of prospective DA is described in Chapter VII. *Excerpting criteria* of this Introduction. The use of regular expression syntax in excerpting filters in text editors.

LNS. List of potential DA devoid of homographs and grammatical codes.

LNS(K). Selection of words that are contextualised on a trial basis. Checking the informational quality of the contexts. Phase II of the corpus queries.

LNS(L). A list of potential DA lumped into base forms (e.g. infinitive for a verb, singular nominative for a noun, etc.).

DA(BVD). List of DA before verification in the Russian language dictionary word collection.

DA(AVD). The DA list depleted words detected in the word set of Russian language dictionaries. Verification of the presence of words was done by automated as well as manual means. Dictionary queries. A description of the verification is described in the next chapter of the Introduction.

DA(BVC). DA list before verification in contextual use. Word validation and first contextual observations.

DA(AVC). List of DA after contextual verification with attached metadata such as newspaper title, number and date. Phase III of corpus queries.

The contextualisation activities and processing of the material that separates the DA(BVC) list from the DA(AVC) list took more than ten years. Obtaining a DA(AVC) list with more than 10,000 rows of contexts made it possible to create the Repository.

V. Verification of dictionary agnonyms

As already mentioned, the acquisition of DA and the processing of each word involved automated activities – using advanced computer tools, primarily a morphological analyser and software for navigating the corpus of electronic texts. Obtaining the 10,000 contexts and the metadata of these contexts in the corpus was done systematically and manually, i.e. without the use of tools allowing, for example, the use of mass (i.e. for a designated number of excerpts), ‘wholesale’, chaotic excerpting of random contexts. The morphological analyser (AOT technology) used for excerpting is integrated with the dictionary *Грамматический словарь русского языка* (Зализняк 1987), or, more precisely, with the grammatical data provided in this dictionary for each word. This is the third⁷ modernised version with 100,000 words. The libraries of the analyser allowed to recognise each word-form of these words also in the CRLPMP corpus. It should be pointed out at this point that the dictionary in question is based on the lexis of four explanatory dictionaries: the oldest, 4-volume Ushakov’s dictionary (1935–1940), the 1-volume Ožegov’s dictionary (1st edition – this is 1947), as well as two academic works: *The Dictionary of the Russian Language*, the so-called MAS, i.e. the 4-volume "Small Academic Dictionary" (1st edition – 1961), and the 17-volume *Dictionary of the Modern Literary Russian Language*, the so-called BAS, i.e. the "Great Academic Dictionary", published in 1950–1965 (cf. Bibliography).

So-called *consultations*, i.e. verifications of the presence of the obtained words in additional lexicographical sources, were carried out. The LNS of words was confronted with the list of words included in the dictionary set – cf. *Verification dictionaries* in the Bibliography.

In this way, the LNS containing (potentially) neologisms was confronted with the grid of words already recorded – before 2006. It should be strongly emphasised here that the present dictionary contains only those words which were not specifically included in the grids of

⁷ There is now a sixth edition of the dictionary. The official online version of the dictionary can be found at <https://gramdict.ru/> – WEB: 22.12.2023. The dictionary has about 110,000 words, including 8,000 proper names.

the above-mentioned sources. Words in the IDA were not verified in dictionaries published after 2006, and data from other corpora that appeared after that year.

Over the course of its contextualisation, the LNS, and ultimately the DA(AVC), has undergone updates, modernisation and depletions, replacements and additions on more than one occasion. The current technical name of this list – is the DA Index in order *a fronte*, i.e. the alphabetical arrangement of entries from left to right.

VI. Context and definition

Before beginning to describe the structure of the dictionary, I would like to make a comment on the essence of the whole issue: why was the procedure of contextualising the LNS words obtained long ago – in 2006 – so important? Well, the procedure in question served first and foremost to verify grammatical correctness and to confirm the rationality of usage: after all, it is the context that makes a word what we call a lexeme. It is secondary, however, whether it will remain in the system of similar discrete entities "from space to space" or undergo a process of agnonymisation, i.e. move from the periphery to a zone which is not reached by the radius of the users' linguistic memory. In the introduction to the sixth edition of his dictionary, Anatoly Zaliznyak writes:

(...) among the hundreds of new words that can be found in the pages of the current press, it is still very difficult to distinguish between those that will remain permanently in the language and those that turn out to be mere one-liners. Moreover, many of these words are not yet stable in terms of phonetic composition, accent and usage. The task of covering this mass of unstable new lexical material should be handled by special dictionaries of new words and meanings, but not by grammatical dictionaries. (Zaliznyak 2022 from the preface to the online version of the dictionary.) [Translated from the Russian – Y.F.].

In order to capture this unstable lexis mentioned by Zaliznyak, it is the context that is needed. In other words, providing the fact that the word has already 'started its career' within the lexicalisation. Moreover, the very fact that a word finds itself in a context already enriches the semantic system: the word enters ready-made semantic fields, updates such fields or initiates new fields as well, makes the gravitation between words more palatable and varied, strengthens syntagmatics, supplements lacunae and linguistic and conceptual minutiae, strengthens the a priori possibilities of language, optimises the user's navigation and orientation system in the extra-linguistic space-time – in other words, *de facto* all that supports and modernises the way of thinking of a modern

language user. Another undeniable fact is that such words – of course, not all of them – actually become lexicalised, settling not only on the closer or further periphery of the *usus*. Some of the new words, due to intellectual demand, may remain only as successful or unsuccessful journalistic occasionalism or brilliant authorial hapax, which, however, does not continue the lexicalisation⁸, does not serve as a prefabrication in subsequent contexts. In addition to the rationality of the use of a new word, the correctness of its use is important; the word is supposed to meet criteria that correspond to generally nomological (i.e. regular) linguistic principles.

The RDA is not prescriptive, not based on normative codification, including orthography. The preserved spelling is original. It can be described as factual or as idiological, i.e. event-based, referring to linguistic facts supported by temporal-local information. It serves primarily to demonstrate the increased activity of the emergence of new words during the turn of two millennia on the territory of modern Russia.

I will stress that it was not my intention to observe the fate of individual words. The RDA does not provide information on how long individual words have existed, whether they have usurped, whether they have been included by dictionaries of new words, etc. It goes without saying that such a powerful – based on an almost ‘unlimited’ textual mass – and lexicographical verification was worthwhile ‘alone’. But "alone" not every research opportunity could be realised. Nevertheless, for the next editions of the RDA, the parameters associated with wider notation are envisaged.

Definiens are also not included in the keyword articles. I believe that definiens are worth formulating for usable neologisms – i.e. those that have become embedded in hundreds, thousands or even millions of contexts and are still occurring, i.e. contextualising themselves. Verification of presence in the current *usus*, observation of the fate of individual headwords, their semanticisation and determination of the algorithm as well as the explicative activities can be a task for other studies.

⁸ When I speak of lexicalisation, I mean a permanent process. Stopping lexicalisation is the non-appearance of a word in a collective context – the so-called *multi-context*, i.e. texts currently generated by individual users, in individual idiolects – the individual languages of the users.

VII. Excerpt criteria

I distinguish four more important criteria for adding words, contexts and metadata to IDA with subsequent inclusion in the Repository:

1. No presence of DA in the collection of *Verification Dictionaries* (cf. the subsection of the same name in the Bibliography chapter);
2. Selection of context with an earlier publication date;
3. Readability of corpus data;
4. The self-sufficiency of the context and the information it contains;

According to criterion (1), the word attached to the IDA should not be present in previous lexicographical sources (cf. Bibliography – *Revision Dictionaries*. For example, I was able to eliminate the following words recorded in other dictionaries:

- Котелова 1995 (*Словарь новых слов русского языка* (середина 50-х – середина 80-х гг.)): *актировка, виброрейка, детприемник, кинонаблюдение, котлоагрегат, культстан, магазин-клуб, рок-музыкант, рок-сцена, телемотрение, телеящик, фильмобаза, фотожурналистка*;
- Котелова 1996 (*Словарные материалы... – 1986*): *гипнабельный*;
- Котелова, Денисенко 1996 (*Словарные материалы... – 1987*): *хард-роковый, шоу-группа*;
- Соловьёв 1996 (*Словарные материалы... – 1989*): *киножитие, киножурналистика, киномонстр, общак, пипл, полузапрещенный, снэк-бар, украинофобия, ультрамарафон*;
- Левашов 1996 (*Словарные материалы... – 1998*): *"химioфобия", самсебяздат, телемагнат, ультраранний, шоу-представление, эмбриотоксический*;
- Буцева 2004 (*Словарные материалы... – 1990*): *абсолютка, эротик-шоу*;
- Милековская et al. 1996 (vol. III): *сверхдоходы, чоппер*;
- Милековская et al. 1990 (vol. II): *фашизонд*;
- Милековская et al. 1988 (vol. I): *кок-чай*.

Of the CRLPMP contexts containing the search word, the context with the earliest possible date of publication in the online version of the local newspaper was selected. This selection was criterion (2). From

among several potential excerpts, I selected the earlier ones, i.e. i. and iv. – see below:

- i. Бодро прошел концерт томской группы "Будни лепрозория". Последние пару лет ребята, "засветившиеся" еще в 80-х, играют с драм-машиной. В их составе давно уже нет саксофона, впрочем, трубы, гитары и баса вполне достаточно для мощного драйва, инициируемого вокалистом и гитаристом Глебом. <Континент Сибирь №7 (249) 2001>.
- ii. После безуспешных попыток найти подходящую ритм-секцию "Шутки" ограничились **драм-машина**. <Псковская губерния №2 (221) 2005>.
- iii. Чаще всего там появлялась с наркотиками подручная "**наркобаронессы**" Эльвира, но могла "осчастливить" жаждущего клиента и другая помощница – Елена. <Любимая Газета – Пенза №38 2004>.
- iv. Клиентами "наркобаронессы" являлись в первую очередь ранее судимые, причем по статье о незаконном обороте наркотиков, жители города Никольска и Никольского района. <Молодой ленинец №20 (6961) 2000>.

Criterion (2) relates to the temporal location of IDA words and is related to the activity of selecting the context created at the temporally earliest point. I recognise that the proposed IDA dating is *de facto* related to the activity of first obtaining the dating (i.e. Dating 1.0) only within the CRPLOM. I also recognise that the clarity and accuracy of the criterion in question may have been "obscured" by a). subjective choice of context, b) the simple absence of other context, c) the lack of dating on the website of an article that originated earlier. Despite these points, I recognise that I have methodically followed the principle implicit in the given criterion – to select the context of use (if, of course, there was such a choice) at the earliest point in time.

But why are IDA modification activities possible and advisable? This concerns first of all the change of dating, i.e. a possible shift (only a "lower" reduction) of the boundaries of Datation 1.0. The first full redaction, i.e. Datation 2.0. can only be a collective work and the next stage of linguo-geochronological verification of the dictionary agnonyms found in the IDA. Thus, the DA would join a common base linking in-

dividual points on the hypothetical dating map of other words as well. Practically speaking, the first anticipated activity during redating is to check the word in other textual resources and to observe the fate of the word – i.e. occurrence in contexts up to a certain point in time from the recorded moment *ab inculnabūlis*. Nevertheless, the already current version of the RDA (with the 1.0 dating carried out in the text space of the CRLPMP corpus alone) can be used for broad linguistic studies, in particular neographic/neological, lexicological, morphological or linguo-chronological studies.

As a separate criterion (3), I distinguish the readability of the corpus data. By this I mean the correct readable encoding of HTTP protocol pages containing newspaper articles. At the time of the creation of the CRPLOM, the most popular encodings were WIN-1251 and KOI8-R, which in certain situations (e.g. in the case of conflicting encodings) made it impossible to read the data correctly. At the time, Unicode encoding (UTF-8) was not the standard for encoding ‘greyed-out’ symbols on the web, as it is today. During excerpting operations, the greatest emphasis was placed on the nature of both the grammatical nature of the word and its unique diacritical properties. It can therefore be concluded that the method used is a diacritical excerpting method, taking into account different types of encoding. One of the technical problems was precisely the disambiguation of the encoding of even entire archives.

The webpage with the newspaper article had to include the date of publication, at least the year or number. Figuring the number of the newspaper (i.e. the print equivalent) on the website made it possible to start a ‘lexicographical investigation’ – to arrive at the correct dating of the excerpt.

The self-sufficiency of the context and its informativeness as a criterion (4) means that out of the possible contexts of DA use, one was selected from which a possible meaning could only be intuitively inferred. The meaning of new words is not entirely stable and fixed. A context consisting of one sentence often had to be expanded even to several sentences or a sentence from a much earlier or later part of the text had to be recalled. In addition, a significant proportion of RDA – are simply words created for the moment. It should also be taken into

account that, when starting to formulate definitions, one would have to juxtapose the meanings of the neonyms with the new meanings that have chosen the existing graphic form. In other words, the IDA does not contain neosemantisms – as, for example, *the Russian-Polish dictionary of lexical innovations* by Bożena Hryniewicz-Adamskich. To give an example: for the entry валюта, the author, in addition to the basic meaning "1. currency (...)" also gives a semantically innovative meaning – "2. *metaphorically* name for the colour of the car coachwork (grey-green)" (Hryniewicz-Adamskich 2005:16). Thus, the neological meaning here is 2., which, however, is not a neonym (at least in its graphic form). I realise that, as part of other research – for example, an analysis of the presence of words in current (in 2024) usage would need to be done in terms of definition. IDA words from the decade 1996–2006 that remained in *usus* would then be called ... *neologisms*. This is because the assignment of neologism status to a word occurs when it is known that a once-new word is still in use. Here the essential word – is "once", for neologisms are "words, meanings of words, idioms that usually exist over a certain period of time (...)" (Kotelova 1978: 22).

It is worth mentioning that the search engine of the press corpus indexer was used in the selection of contexts. The indexer is a programme that made it possible to index, i.e. sort newspaper articles in the archives, and the words themselves according to a predetermined alphabet (cf. below). This succeeded in obtaining discrete linguistic entities – words as segments of alphabet characters from space to space or punctuation mark. In addition to the 66 Russian letters, inverted commas, a hyphen and 52 letters of the Latin alphabet were also taken as alphabet characters. Such unique discrete autonomous graphic entities as "sms-ка", "sms", "SMS", "SMS", "SMS "ка", "смс"-ка, "SMSка" were thus extracted, "смс-ка", "смс", "СМС", "SMSка, смс-ка, смс, СМС, "SMS-сообщение", "sms-сообщение", "СМС-сообщение", "смс-сообщение", "SMS"-сообщение, "sms"-сообщение, "СМС"-сообщение, "смс"-сообщение, SMS-сообщение, sms-сообщение, СМС-сообщение, смс-сообщение – and even "эсмэска" and эсмэска. What is undoubtedly conspicuous is the chaotic nature, the lack of a conventionalised sequence of characters to denote an *sms* text message. A manifestation of the very

novelty of the expression was the use of inverted commas – a sign of the hesitation of the journalists' decision within the use or non-use of inverted commas. It should be mentioned here that two graphic images are currently accepted – *смс* written in the Russian alphabet and *sms*. In my book titled *The Excerpting of Quotation Marks from Russian Press Texts* [Fedorushkov 2009], I mention that inverted commas have an adaptive function during the fixation of a word in the usus before the neologism (or more precisely: its sound, its acoustic image) adopts a fixed graphic image – a conventionalised standard in grammar and orthography, i.e. it becomes embedded in the linguistic norm. The inverted commas in the neologism during usualisation mostly disappear, and the initial or 'transitional' function of the inverted commas is adaptive, i.e. metatextual. Therefore, examples of type (i) and (ii) were attached to IDA:

i) "**авиализинг**"

На это ему посоветовали поработать по вхождению завода в бюджет по строке "**авиализинг**" и за неделю снизить долги по зарплате до двух месяцев. <Неделя области № 43 (108) 2004>.

ii) **авиализинг**

Активно сотрудничаем с лизинговыми компаниями России. На сегодняшний день **авиализинг** – спасательный круг отечественного авиапрома. <Капитал-Weekly №14 (513) 2005>

Why was the selected period 1996–2006 ? Why weren't more symmetrical, *strict* dates taken into account, i.e. e.g. *strictly* the decade – 1995–2005? Well, the work on the LNS took place precisely in 2006 and 2007 with the intention of analysing the preceding 10 years, i.e. 1996–2006/2007. Therefore, I have decided to call this period "decade", although it is actually eleven years in length. Furthermore, in the first half of the 1990s, there were simply few online versions of local, i.e. regional, newspapers. This will be discussed in the next chapter.

VIII. Historical period: the decade for the "language of newspapers"

The period in which the discussion around the notion of *newspaper language* developed was in the 1970s–80s (cf. Novikova 2002: 38). What is the essence and boundary of this concept, and is there a ‘newspaper language’ at all? Opinions on this subject are ambiguous. Some authors focus on the following view of the phenomenon in question: the language of a newspaper is referred to within the framework of *the journalistic style* (cf. Rogova 1975; Panov 1988) or within the framework of the so-called *journalistic sub-style* (cf. Kozhina 1977: 179). Some authors speak of newspaper language as a functional-stylistic category of language – i.e. in a broader sense than *style* as such (cf. Kostomarov 1971: 257, Rozental’ 1977: 39). V.A. Alekseev questions the existence of newspaper *style* and – more – the newspaper variety of journalistic style in contemporary Russian (cf. Alekseev 1973: 9). This author mentions, among other things: "1. newspaper style and journalistic style; 2. functional-stylistic category, broader than style in the sense of point 1; 3. newspaper version of journalistic style" (Alekseev 1973: 9). Nevertheless, the above and similar works emphasise the peculiarity of newspaper language and the necessity of discussing ‘newspaper language’. An important scholarly opinion in this regard belongs to the researcher V.G. Kostomarov, who argued for the existence of the unity of the language of the press on the basis of the principle that the systemicity of the functioning of style is determined by the unity of linguistic and extralinguistic coefficients and the mutual relation of the units of the different levels of language on the basis of their common communicative task (cf. Kostomarov 1971). Speaking of the rational and emotional impact on the reader, Kostomarov takes as his starting point the principle of the ‘interchangeable occurrence of expression and standard’ (Russ. *чередования экспрессии и стандарта*) (Kostomarov 1971: 85), which ensures „the reliable transmission to the reader of the factual and informational side with its commonness, effectiveness, rigour, unambiguity, intellectual relevance, and the influential and organisational side with its exclusiveness, emotionality, even sensatio-

nalism" (Kostomarov 1971: 85) [translation – Y.F.]. O.V. Novikova believes that to a large extent this principle is explained „from a historical point of view" (cf. Novikova 2002: 38). Thus, for example, researcher N.L. Sinyachkina argues that it was the centralised system of authority (including – the requirement of accuracy for messages revolving in the orbit of the main authority) in the 1970s–80s that proved to influence the rotation and use of lexis. Allegedly beginning in the 1920s, this system influenced the formation of a stylistic canon and stylistic clichés in the language of the newspaper (cf. Sinyachkina 1997: 8).

Before the collapse of the USSR, during the *perestroika* period, colloquial style was not allowed in the language of newspapers, which was dominated by bureaucratic style. From 1986 onwards, the standardisation of lexis in Russian began (cf. Novikova 2002: 38) and with it the disruption of literary norms. It seems necessary at this point to detail precisely the historical moments in the evolutionary changes that occurred in both the national language and the language of the press.

In the last decade, IT online resources have been creating strong competition for the printed press in Russia. The popularity of online information publishing has grown rapidly, so that now virtually every regional print periodical in the Russian Federation has an online representation. A significant influence on this state of affairs was the liberal policy initiated in 1985 by Mikhail Sergeevich Gorbachev (cf. Chubinsky 1998: 474–493, Vepreva 2005: 11–15), which laid a solid foundation for a change in the nature of information broadcasting in the media, especially in the press. The advancement of the transformation meant that Gorbachev's team began to change its attitude towards the West. As a result, among other things, the jamming of some Western radio stations broadcasting programmes about the Soviet Union ceased in 1987, and Western newspapers began to appear in cities (cf. Czubinski 1997: 375). At the same time, the media were advised to abandon their previous monologue forms and introduce dialogue in their programmes, and, moreover, to cover social topics more frequently, but in a way that excluded mentoring and lecturing (cf. Gorbachev 1989: 95–99). The given state of affairs was termed *glasnost*. As a result, the Soviet Union became 'the most read country in the world' (Donaj 2001: 33). "Loosened" and then statutorily abolished censorship in August 1990

and the introduction of freedom of speech resulted in the appearance of content in the press that was independent of previous party-state prescriptions (cf. *ibid.*).

Democratisation in the social, political, economic sphere meant that some of the newspapers that were organs of the party or the state were privatised and passed into the ownership of the editorial board (and in practice: passed into the hands of so-called private capital), which – through the editors-in-chief – decided on the form and content of the articles themselves (e.g. in the newspapers *Комсомольская правда*, *Известия*, etc.); extra-central ‘suggestions’ (i.e., e.g., arising from orders or bans directed towards the editorial office by the commercial owner of the media) are omitted here due to the lack of documentation on this subject. Other periodicals, on the other hand, which had a circulation of millions, lost significantly in readership numbers after the collapse of the USSR: for example, the newspaper *Труд*, which had a circulation of 21 million copies in the period 1986–1991, owned by the editorial board, had a circulation of 1.3 million compared to the Moscow weekly *Аргументы и Факты* (the so-called *АиФ*). The circulation of this newspaper at the time was 32 million copies sold throughout the USSR. This was the highest single circulation in the world (cf. Donaj 2001: 39-40). After the collapse of the USSR, the weekly newspaper *АиФ* recorded a circulation of 3,300,000 copies. In conclusion, it can be said that press liberalisation was the earliest and widest form of media transformation.

The printed press found itself in opposition to the ideologically oriented television and radio media (cf. Drawicz 1997: 48-62). However, in recent years, the demand for the printed press has decreased significantly (this is more of a worldwide trend). Russians no longer read as much of the printed press as they once did, at the time of, for example, the systemic reconstruction in the early 1990s. This is related to the pace of life, the technological rush and the increasing number of people with access to the internet, which provides a convenient way to access a wide variety of press items, e.g. the content of a new issue and archive information. It is worth noting at this point that some newspapers make electronic versions of a new issue available in full only some time after the print version has appeared.

Access to the archives and, above all, the electronic archives of the Russian press before the advent of the internet was difficult. The internet ‘existed’ in Russia from 1990, with the ru domain registered for the Russian Federation (example: www.yandex.ru) only coming into existence in 1994. This was the beginning of the globalisation of the internet in Russia. In 1995, in turn, the www.ug.ru newspaper *Учительская газета* was opened. This event triggered the beginning of the development of digitalisation and digitisation of printed periodicals with the provision of data on the Internet. This newspaper, perhaps, became an example for other Russian periodicals. An online newspaper *Газета.ру* without a print counterpart was also created.

G.N. Trofimova defines contemporary network publishing, especially mass media (Russian abbreviation *СМИ*) as follows: *электронные, сетевые, онлайн-овые* and *виртуальные* (cf. Trofimova 2001b).

In this paper, I only consider material located online (i.e. in electronic form), and therefore I am interested in material that – in electronic form – has a printed counterpart. With a view to shifting the centre of gravity to the online edition or getting rid of the print publication altogether, I must also point out the fact that some content appeared exclusively online (in the electronic version of the newspaper) back in the day. Periodicals that existed in the form of ‘print version = online version’ were not the only ones included when the CRPLOM was collected. However, I am aware of possible exceptions to this situation. I do not consider this a shortcoming that would significantly affect the idea of excerpting ‘print-only’ material: after all, readers have the right to choose, and electronic and print content began to function in Russia as competitors to each other.

Prior to the advent of the Internet in Russia, the possibility of collecting and studying the press vocabulary of each Russian subject was associated with the simultaneous work of many researchers. The hitherto unprecedented possibility of linguistic observation of Russian language vocabulary extracted from the electronic press has been exploited by various scientific groups, including the *Integrum* company⁹ and the creators of the¹⁰ project *Компьютерный корпус текстов русских*

⁹ Cf. website <http://svc.integrum.ru> – WEB: 12.12.2007).

¹⁰ The corpus and the online version of the corpus were created at the Laboratory of General and Computational Lexicography at the Faculty of Philology, M. Lomonosov University, Moscow. The author of the project – A.A. Polikarpov.

газет конца XX-ого века. The creators of the project used 13 complete archives from 1994 to 1997 of such newspapers as *Завтра*, *Правда*, *Литературная газета*, *Новая газета* and others.

In the present study, the writer of these words also took the opportunity to linguistically observe the vocabulary of Russian in the electronic resources provided, with the selection of multi-genre newspaper material based on a search and selection of newspaper information in individual RFSs. The expansive growth in the number of online publications is reflected in the number of texts, and therefore in the popularity of the journalistic sociolect, so also in words – including neonyms. I visualise the increase in the number of neologisms in the FR local press in a series of 11 infographics for 11 years. The intensification of the colour scheme (from yellow to purple) in the individual RFS signifies precisely the expansion of DA temporally and geographically:



Figure 3. Occurrence of DA in the local press in different RFSs in 1996 [own study]



Figure 4. Occurrence of DA in the local press in different RFs in 1997 [own study]



Figure 5. Occurrence of DA in the local press in different RFs in 1998 [own study]

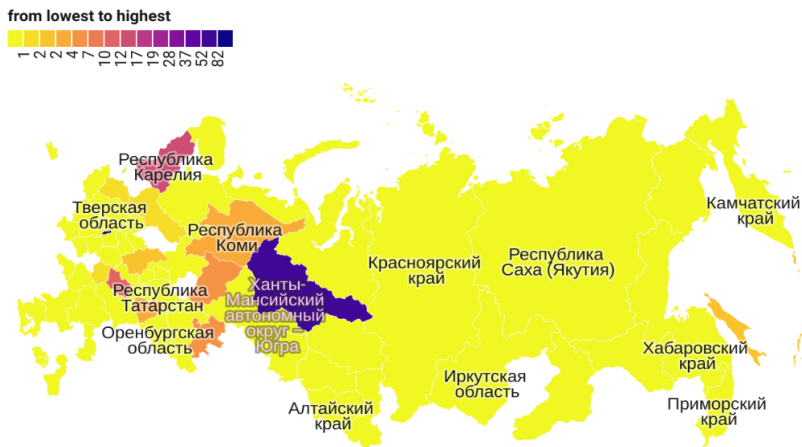


Figure 6. Occurrence of DA in the local press in different RFs in 1999 [own study]



Figure 7. Occurrence of DA in the local press in different RFs in 2000 [own study]

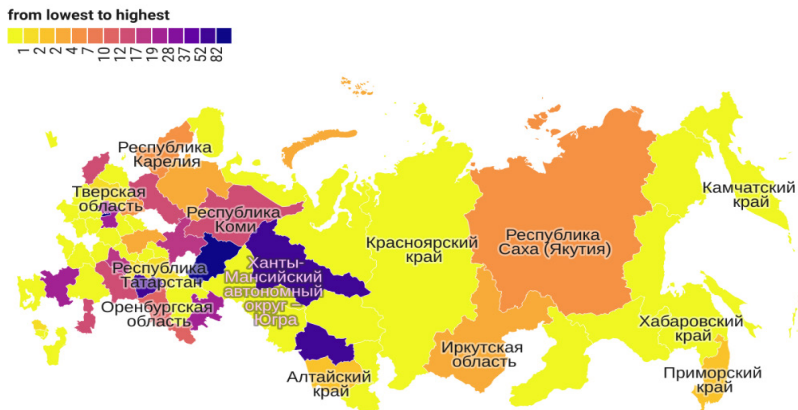


Figure 8. Occurrence of DA in the local press in different RFSs in 2001 [own study]

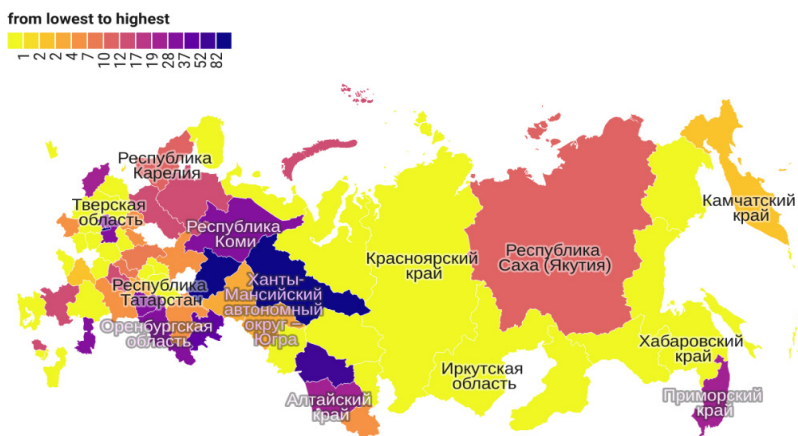


Figure 9. Occurrence of DA in the local press in different RFSs in 2002 [own study]

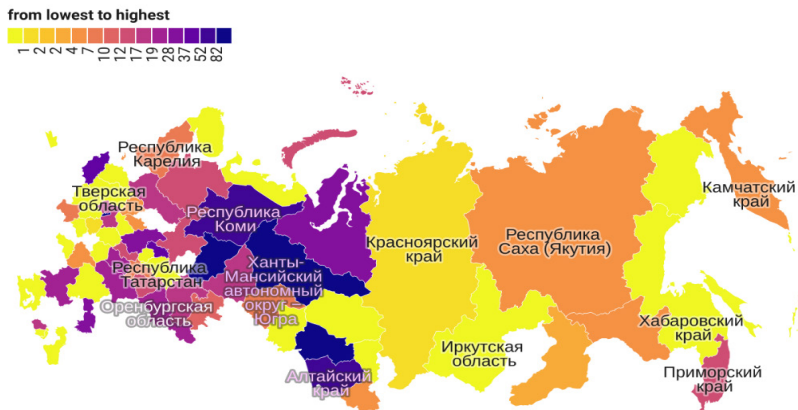


Figure 10. Occurrence of DA in the local press in different RFSs in 2003 [own study]

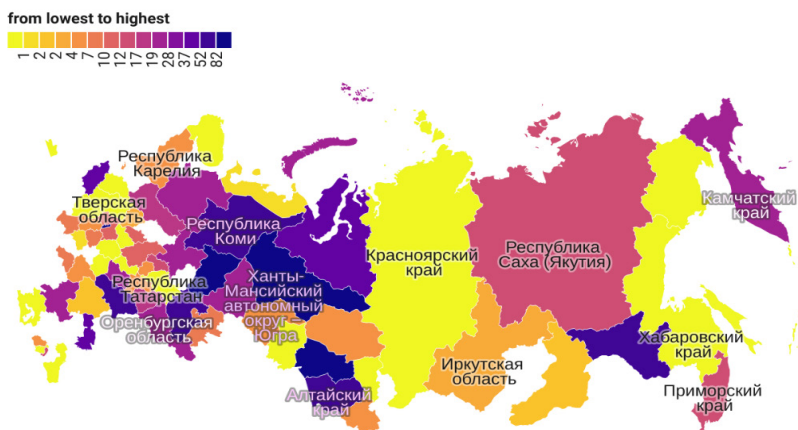


Figure 11. Occurrence of DA in the local press in different RFSs in 1994 [own study]

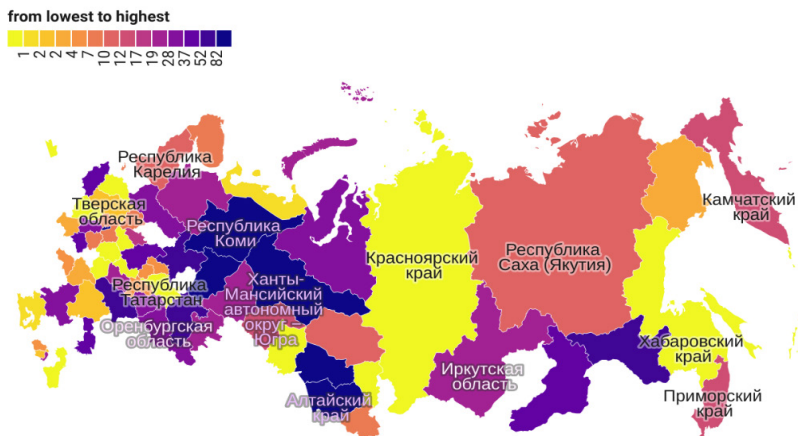


Figure 12. Occurrence of DA in the local press in different RFSs in 2005 [own study]

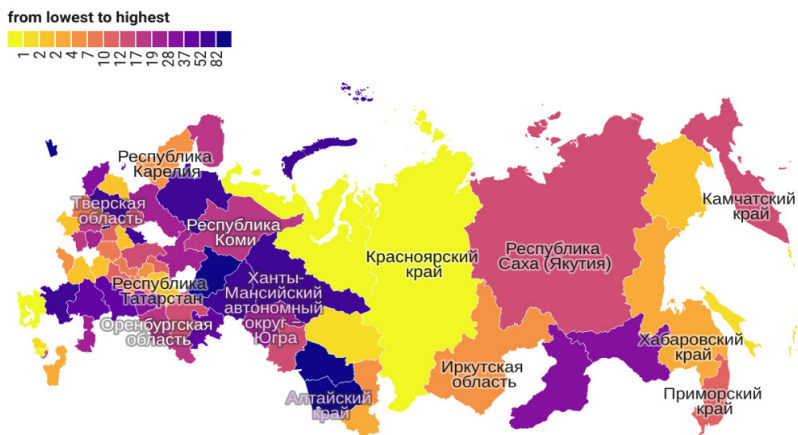


Figure 13. Occurrence of DA in the local press in different RFSs in 2006 [own study]

From the above infographics, it can be inferred that the growth of neonyms over the millennium decade correlates with the West-East line, with the peak of the explosion (for most regions) being 2004–2005. The primary factor in the growth of DA – primarily neologisms – appears to be the emergence of newspapers themselves in web format, although it can also be surmised that the growth is related to the following factors related to the public motivation to create an online version:

- optimisation of the economy and positive changes to regional funding;
- the expansion of Gazprom and other corporations into hard-to-reach areas;
- dissemination and stabilisation of the internet network;
- freedom of expression of private ideas and political opinions;
- access to non-Russian media (press, television);
- optimisation of the quality of rail services and other transport systems, including – optimisation of transport infrastructure, construction of federal routes;
- increased revenue, public welfare and a wealthier society with opportunities for public purchasing;
- intensive development of communication and computing technology;
- positive pro-social political developments;
- opening up to the West and Western technology;
- the possibility for citizens to travel the world;
- "bringing" the internet to individual regions, i.e. developing infrastructure for connectivity in general.

The current version of the RDA lexicographical resource makes it possible to make a preliminary research forecast: DA over the period under study tends to develop towards subject areas such as technology (including hi-tech, computer), IT, business and marketing, economy, music culture. This is presumably linked to Russia's wide-ranging opening to the West after the collapse of the Soviet Union. The media were advised to abandon their previous monologue forms and introduce dialogic speech into the discourse, addressing social issues but in a way that eliminated mentoring and lecturing; the liberalisation of the press was the earliest and broadest nationwide form of media transformation. This is confirmed by the researcher Kostomarov's accurate

statement: ‘(...) the literary and linguistic norm becomes less definitive and obligatory. Russian language use is now characterised by the main orientation of modernity towards freedom, with rampant self-will, with carelessness – the term *liberalisation* is appropriate to describe these very rapidly developing processes’ (Kostomarov 1999: 30) [translation – Y.F.]. However, an analysis of the thematic development of the vocabulary requires separate research.

IX. Research opportunities in morphology

Therefore, the Repository can be considered a representative and useful source for researchers in the field of morphology of vocabulary. Speaking about the trends in the development of morphological structures, it should be stated that the vocabulary actually signals "out of control" of the norm and censorship in spelling: *B2B-маркетолог, sms-голосование, sex-шоп, free-чек*), graphic and morphological structure (*слэмить, СПИД-"УАЗик", CD/MP3-плеер*), thematic and stylistic (*спиммер, порноюмор, поурбаниться*). Affixal formants from foreign languages – primarily English – play a special role here. And although the morphological aspect is not within the research scope of the present project, an intensive increase in the use of borrowed affixes and radices with technological themes, i.e. related to the rise of modern technology (e.g. *веб-устройство, pdf-файл, флэш-карта, биоэнерготерапевт*), is observable in the period 1999–2003. From the first observations and preliminary counts, it is clear that in terms of morphological structure, the neologisms of the millennium period show exceptionally analytic, agglutinative features. Initial formants (prefixes and prefixoids) from foreign languages – primarily English – play a special role here.

The researcher Joanna Misturska-Bojanowska, in the summary of her monograph entitled *Динамические процессы в словообразовании имен существительных в русском языке 2-й половины XX века* (Eng. *Dynamic processes in the word formation of nouns in the Russian language of the 2nd half of the XX century*), notes: "Our analysis shows that the following processes are most active in the sphere of noun production in the second half of the 20th century: internationalisation, the dynamics of quantitative growth of lexemes with personal meaning, the extensive use of proper name as a word formation base, the strengthening of compression in word-formation, colloquialisation and the growth of expressivity, the growth of nominal prefixation, the competition and specialisation of certain affixes (suffixes and prefixes), the spread of synonymy, antonymy and word-formation homonymy, trans-step word-formation, the growth of analyticism and agglutinative features" (Misturska-Bojanowska 2013: 137–138). Some of the afore-

mentioned processes can be seen ‘with the naked eye’ (i.e. without detailed statistical examination) in RDA.

Czech linguist František Trávníček distinguished between three types of hybrid: "(a) the word is formed from a native or assimilated word-base by means of a foreign suffix, (b) a foreign prefix is added to the native or assimilated word base, (c) in a foreign compound word, one part is replaced by a native one" (Trávníček 1948: 30–31 after Mielta 2001: 52).

It is no accident that I cite this peculiar classification. But it turns out that the given classification (described, after all, for Czech, in 1948) is also valid for the tendency of the development of Russian word constructions observed (tentatively) in RDA. In other words, RDA can be used in confrontational studies, e.g. trans-Slavic combined with a chronological parameter. I also suppose that practically any classification concerning the vocabulary trends of the recent period in Russia (after the collapse of the USSR in 1991) can hypothetically be (and is worth) confirming, based on the RDA resource.

I extend the terms prefix, prefixoid (and suffix, suffixoid) within the limits of the study to the term *word-formant*. Examples of initial formants:

sms-, авто-, азото-, айф-, аква-, анти-, аудио-, ауто-, афро-,
аэро-, байк-, байто-, баксо-, би-, биг-, бизнес-, билайно-
био-, биотехно-, бифи-, блюдо-, бого-, бомбо-, брейк-, бэк-,
валюто-, веб-, вебо-, вело-, взаимо-, вибро-, видео-, випо-,
власте, вне-, внутри-, воен-, второ-, вузо-, вшиво-, высоко-,
газо-, газонефте-, гейм-, геохимо-, гетеро-, гидро-, гимно-,
гип-, гипер-, гипо-, глав-, годо-, гомо-, гос-, гранд-, дам-,
дарт-, дач-, де-, дем-, джем-, дип-, диск-, до-, долго-, дон-,
дор-, дорого-, дура-, дух-, дымо-, евро-, едрос-, за-, зиц-,
идео-, интернет-, инфо-, исламо-, ист-, камне-, кантрик-,
квази-, кверху-, кино-, киновидео-, кисло-, книго-, комиксо-,
коммуно-, консерво-, контер-, контр-, крано-, кригс-, кубер-,
кубино-, кувалдо-, культууро-, курорто-, кучма-, лав-, ланд-,
ландро-, ласси-, латино-лево-, лега-, легко-, ледо-, лейко-,
лесо-, лже-, ливне-, лизо-, ликеро-, лио-, листо-, лит-, литро-,
ложно-, лозо-, лонг-, лонго-, луго-, лыже-, мавро-, майстха,
макс-, макуло-, мало-, мас-, машино-, мега-, мед-, меди-,

медиа-, меж-между-, мезо-, мент-, место-, мета-, метод-, метро, мех-, мигранто-, микро-, микровидео-, мимо-, мин-, мини-, мино-, мифо-, младо-, много-, моб-, моно-, мор-, москво-, мото-, мотоцикло-, мульт-, мульти-, мульто-, на-, найто-, нано-, нарко-, не-, недо-, нейро-, нео-, нефте-, ново-, ноль-, о-, об-, обл-, обмано-, оген-, одно-, орг-, ортодром-, орф-, осексо-, оф-, пара-, парт-, плей-, пневмо-, поли-, полно-, пост-, про, проф-, пси-, психо-, психоневро-, радио-, ретро-рок-, рэп-, сверх-, скреб-, смехо-, смс-, софт-, спеле-, стар-, супер-, теле, телевидео-, телевидеофото-, термо-, тетра-, техно-, фильмо-, фито-, флай-, фото-, фри-, харт-, хэнд-, черно-, чечено-, чисто-, чрез-, шаро-, шело-, шеф-, шино-, шир-, широко-, шифр-, шопо-, шорт-, шот-, шоу-, штраф-, шутц-, щедро-, щито, эго-, экзо-, эко-, эколесотехно-, эконом-, экс-, экстра-, экстрим-, электро-, эмбо-, эмбрио-, эндо-, эндовидео-, энерго-, энио-, энтеро-, энто-, энхазл-, эпид-, эро-, эст-, этно-, эхо-, южно-, ява-, языко-, японо- etc.

Only a preliminary observation of the IDA shows that the number of borrowed affixes and affixoids exceeds a thousand. Presumably, many of these are neonymic in nature. The present project in its realised form offers research perspectives within the analysis of morphological and morphotactic development. The issue of morphological properties – especially within the clarified classification "affix – affixoid – semi-affix – radixoid – radix (stem)" (Minina, Bartkov 1980, Bartkov 2006, Bartkov, Minina 2019) – may be a continuation of the research described in this project.

Therefore, it should be concluded that RDA can be used for broad linguistic research, in particular lexicological, morphological, neological research. The integrated information generated in RDA can also serve as a basis for research in language history, sociolinguistics, journalism, lexicology, lexicography, media studies. The significant significance of the project in relation to previous research results on the unique vocabulary of the millennium period will be that the study will result in the world's largest repository of previously unrecorded words in Russian.

X. The discreteness of the word and the inadequacies of the RDA

A neologism – and this is the central research object in this description – represents not only a new acoustic or graphic image. It is, after all, linked to a mental act, leaving behind a new imprint on the map of mental images correlated with designators – names of situations, phenomena, processes, objects, features, states. In contrast, any verbalisation, i.e. the creation of an utterance, is a locutionary act which, compared to a mental act, is a secondary precedent. It is a finished image, beginning its evolution in the contexts of use of individuals, i.e. in individual languages, also called idiolects. A multilect – i.e. an area in which contexts of use accumulate – is then called a collection of individual idiolects. Here we can compare the accumulating contexts of textual use with culture – the shared identification, recognisability, codification and connotation of a new discrete entity – the neonym, i.e. a unique unregistered rationally used combination of alphabets/ alphabet letters with an existing meaning (neonym-‘neograph’) or a new, unregistered meaning (neonym-‘neograph’ – neosymantism, i.e. simply neologism). According to anthropologist Adam Czyżewski, a new image – including a neologism – understood as a precedent, first of all creates a newly created convention, i.e. initiates a culture of use. In turn, the generation of a context of use – in idiolects – is also something secondary. Czyżewski points out: "The inability to recognise an image is due (...) not to the lack of an adequate context, i.e. socially transmitted perception schemes, perceptual patterns, functioning in the mind of the native surprised by a completely new form of visual message, but simply to the lack of precedent" (Czyżewski 2009: 30). A neologism also represents a point of reference that requires a specific moment in which idiolects adapt to the cultural impulse, to the type of locution and at the same time to the conventions of use. Such a moment Adam Czyżewski calls the "critical moment of fusion of idiolects" around itself, when the effect of the birth of a community of "reception" and "recognition" (cf. *ibid.*) of a new acoustic/graphic image – i.e. a neologism – manifests itself. A neologism understood as a new image is

a precedent *per se*, and has an occasional, peripheral and agnonymic status. This status enables some agnonyms to make their way from the periphery to the active and current centre of lexical usus in its most vigorous and energetic realisation – use and duplication in currently emerging *hit et nunc* texts. The neologism passes into other idiolects, i.e. it figures – like a quantum – simultaneously in several locations – at several points in the multilect that is language as understood sociolinguistically. However, it should be noted here, citing Czyżewski’s opinion, that the concept of culture encompasses any ‘arbitrary, individual mental act which, in the process of perception of images, precedes the moment when the pressure of the context, most generally understood as knowledge, is revealed’ (ibid.). The birth of a neologism in an idiolect is the result of a unitary mental act, and the ‘demand in birth’ of such a neologism is theoretically multielectal. To a certain extent, such a birth is predictable if such a new entity actually fits into a ‘contemporary’ context, revealing by contrast with other entities (words) its precedent. It can also be said that a neologism, while beginning to duplicate itself in a multicontext, has something in its meaning (i.e. a cluster of semes – a semanteme) that sets it apart from the meanings of other words. It can conventionally be called the suprasememe (i.e. the highly implicit ‘over’-sememe) of precedence. Hypothetically, contexts that absorb such a meaning precipitate such a supraseme over time, so the neologism loses its most salient feature – *novum*. Predictable, then, is also the path of dissemination, usualisation, so that the neologism also adopts a certain structure (mental → phonetic → morphological, semantic → valency) against others (earlier, time-identical, potential). The predicted emergence of neologisms can thus have the qualities of a mathematically sound prediction. In other words, it is possible to predict what morphological structure such a new acoustic and/or graphic image will have. If, on the other hand, we have a solution to the problem of determining the state of culture and the mental background in general, to the demand of which language also responds by means of neologisms, periphrases etc., then the prediction of the appearance of a neologism in speech is a theoretically possible thing. This is due to another parameter of the word – discontextualisation, i.e. the basic attribute of the word, which is related to autonomy in the presence of

other words, i.e. other – to a greater or lesser extent – discrete entities. Contextualisation makes the word a *lexeme*. What follows is a phase of lexicalisation, which is one of metamorphosis, driving and modifying not only the *usus* itself, but also the morpho-composition, affecting the valency potential of individual discrete entities, increasing the masses of new sentences and texts, extremely intense changes in the centre and periphery of the *usus*. Contextualisation provides an opportunity for the continued life of an expression – i.e. its multicontextual rotation. A word is a morphologically and semantically discrete entity, but the law of the economy of language means that semantically a word need not be discrete at all – e.g. in the idiom *black sheep* we are dealing neither with the denotate *sheep* nor with the name of the colour of the fur. The two words form a single super-meaning (the so-called *global* meaning), which takes on the function of discreteness among other discrete entities in the text. Therefore, neologisms should be observed and verified in terms of their discreteness or lack thereof, as well as in terms of their co-occurrence with neighbouring words and, more precisely, their meanings. Moreover, a neologism, having a new graphic-acoustic image, may not occur alone – but always with a ‘neighbour’. This is why the RDA has a fundamental flaw – the failure to take into account the parameter of semantic discreteness: it includes neither collocations nor idioms with such a neonymic ‘neighbour’. For example, the collocation лыжероллерный круг with ДА.лыжероллерный in the article:

лыжеролерный

Сегодня гордость стадиона – это освещенная лыжная трасса и наполовину построенный **лыжероллерный** круг. <Архангельск №(47) 2006>.

Other possible collocations are *попасть в лонглист [конкурса]*, *попасть в шортлист [конкурса]*, *войти в лонг-лист*, *лонг-лист этого года*, *дебютный лонгплей*, *онлайн-музыкальный магазин*, *релиз DVD*, *дебютный лонг-плей*. The result: next to the neologism there may be another word "likened to it". Such a cluster of words may represent a potentially new *phrasemes* in which one of the components is a neonym. Contextualisation makes it possible to search for phrasemes and potential phrasemes (PPs), but also *new* phrasemes

(NPs), i.e. word combinations that together enter the path of phrasematization. One single neologism creates its own ecosystem contributing to something else new. It is possible that the phenomenon of coexistence (symbiosis) of two networks/two systems is at work here. This can be compared to mushrooms and *mycelium*: mushroom pickers know that if you see a fly amanita, you should look for another Amanita m. nearby or... a cep (and vice versa). In a linguistic text, the principle of the coexistence of two semiotic networks may also operate: if you have found a neologism, look for another neologism or something similarly unusual or rare in the immediate contextual surroundings, e.g. unique in terms of grapheme (neonym), unusual in terms of meaning (neosemantism, neologism), a connotation constituting a precedent (neologism), something unusual, escaping the conventions of style, spelling, pragmatics, connotation, etc. With all this in mind, I conjecture that a high number of contextualised uniques in a particular textual set makes it possible to make a statistical prediction of the number of other (or the same) uniques in another similar set. Accounting for contextual parameters is possible with neural networks. In other words, the network can be 'tuned' to automatically search for 'new/unique' – bypassing manual or automatic excerpting. The context coefficient of the co-occurrence of neonyms can theoretically become a vector feature of a neural network. Therefore, I consider the context of occurrence as the most important parameter for dictionary agnonyms. In this connection, I refer to the RDA as a contextual repository rather than a datational or geographical. The contextualisation of a word is also important because context largely makes it possible to figure out what the thematic affiliation of a word is, or even what the meaning of a lexeme is. It helps to realise the way in which a word comes into being for a particular context, i.e. by already overcoming the first barrier towards possible occasionalisation – the transition from the zone of non-being to the zone of existence alongside other lexical units. When generating or regenerating such an acoustic or graphic image, the journalist, other word creator or the person repeating the newly heard word does not distort it, envisaging a common, universally adaptable name for the hitherto unfamiliar or poorly-lexicalised graphic-acoustic referent of the potential name. This applies, for example, to the neologism *евросантехника*:

евросантехника

Встретили нас у поезда, привезли в санаторий, водитель проводил к администратору, и за считанные минуты мы были в номере. Импортная мебель, телевизор, холодильник, телефон, **евросантехника**, пластиковые окна, паркетный пол, покрытый коврами, чистота и порядок. <Контакт №11(423) 1999>.

The degree of ‘relevance’, the actuality of the use of contextual occasionalism signals at the same time a possible prediction of further use. A word-concept gestalt is then created, a code, a convention for future uses in the multilect, i.e. a set of texts generated in a set of individual lects – idiolects. The repeated word reflects a kind of ‘solidification’, routinisation (Bierwiazzonek 2006: 448), as a prefix with a positive connotation is used: the name of a product, good or service with a *euro* prefix – especially in marketing – is correlated with the connotation of ‘something solid, exemplary, noble, worth imitating’. This reproduces the model of the stereotype (Bartminski 1980: 13–14) within fashion, modernity, while the product, good or service is solid and of good quality: e.g. *евровложение, евроремонт, евроотделка*. Although, in addition to the positive connotations (Kobozeva 2000), one can also distinguish between the positive connotations: *еврозависимый, еврозаятка, евробюрократия*. It is interesting to note that the string *евросантехник* actually began to function actively during the period in question; a proper name also appeared, used by specialist companies. The acoustic image of the denotative name routinises faster in speech compared to the fluctuations of the graphic notation (cf. the above examples with the word *sms*). I suppose, therefore, that phonological lability of such an image, i.e. variants in pronunciation, sound, may occur less frequently. However, when widely available mobile phones appeared on the market, the name of the referent in question and the spread of this communication tool fluctuated within the limits of the synonymic order in a no less labile stylistic spectrum with a strong tilt towards the colloquial style: *сотый, сотик, сотовый, сотовый телефон, труба, трубка, мобила, телефон, мобильный, мобильник, мобильный телефон*. Other researchers also mention *сочка, сотовик* (Мандрикова, Кротова 2013: 37). From the given order, only

сотовый, сотовый телефон, мобильный телефон found themselves in the neutral style range with the later predominance of the use of *мобильный телефон*. The word *мобильник* has crossed the boundaries of colloquial style and can be considered a neutral word alongside the modern *смартфон*. However, the repository does not offer the possibility to observe thematic words. The introduction of a thematic parameter is possible in its next editions of the RDA. The current parameters are detailed in columns 1.–9. of Table 4.

XI. Structure of the RDA and access to the resource

Access to the RDA is remote – via the website <https://jefed-repository.web.amu.edu.pl/> (username: *user*; password: *user*).

Navigating through the RDA is possible vertically and horizontally. The data obtained in this way allows the reading of the entire headword article (horizontally), as well as entire sets of words with a specific formant (vertically), while moving through the entire dictionary from the same word in the index *a tergo* to the same word in the index *a fronte*.

Table 4. RDA columns in vertical (columns 1., 2.) and horizontal (3.–9.) dimensions

Vertical dimension RDA		Horizontal dimension of the RDA						
1	2	3	4	5	6	7	8	9
Index a tergo: lemma without quotation mark	Index a fronte: lemma without quotation mark	Index a fronte: original lemma with quotation mark	Context with word form	SOURCE	YEAR	RFS	CITY	MAP

In what follows, I will discuss each column separately. So let's take a look at the different levels and parameters of consultation in the RDA.

XI.1. The vertical dimension of the RDA

Let's look at Table 5. The vertical dimension is the dimension of moving between the indexes in column 1 and column 2. The index in column 1. is a list of words in the *a fronte* IDA layout, the index in column 2. is the same in a different layout – *a tergo*, i.e. in the usual alphabetical layout allowing one to search for words beginning with a particular letter or sequence of letters in alphabetical order from left to right. The decision to include the *a tergo* index as the first in a row of columns arose during a series of public presentations of the RDA attended by a number of linguists specialising in morphology and word formation. This list allows words to be searched from the end of words (i.e. from right to left), which is particularly important for those searching for groups of words with specific grammatical properties, with specific formants. The graphic forms of the two indexes are the same, in contrast to the index in the third column, where the spelling is not without inverted commas, preventing alphabetical ordering of the data.

Column 1. Index *a tergo*: lemma without quotation mark (initial and terminal)

In a given column, the IDA is placed in inversion *a tergo* in three orders: 1) numerical from 1 to 9, 2) alphabetical Latin from *A(a)* to *Z(z)* and alphabetical Cyrillic from *A(a)* до *Я(я)* in the direction from left to right of the character string (word). Such a word is a lemma devoid of paired inverted commas (initial and medial) except for the inverted comma placed in the middle of the word (medial). The alignment in index formatting is from the right. This technical procedure makes it possible and easier to search for words from the end of a word, especially when trying to select words with the same terminal formant. Each word in a given index is linked to the same word in the next column in which the index *a fronte* is placed (cf. below).

Column 2. Index a fronte: lemma without quotation mark (initial and terminal)

The words in a given index are technically related to the words in the a tergo index in column 1. Each DA in column 2. (*a fronte*) is the same as the word in column 3. The difference between the DA in the second and third columns of a fronte is the missing inverted commas.

Table 5. A tergo and a fronte indexes with two-sided hyperlinks – columns 1. and 2.

1	2
Index a fronte: lemma without quotation mark	Index a fronte: lemma without quotation mark
(...)	(...)
лонгплей	лонглист
лонг-плей	лонг-лист
лонг-листер	лонг-листер
лонглист	лонгплей
лонг-лист	лонг-плей
(...)	(...)

I should emphasise that the first two indexes (columns 1. and column 2.) are the same words in a different alphabetical arrangement; they are correlated with each other by a system of bidirectional hyperlinks. This means that the RDA user, by clicking on any word in the index a tergo, will be redirected to the same word in column 2. – in index a fronte. And vice versa. The first two indexes in the first two columns (columns 1. and 2.) are working, facilitating vertical navigation through the RDA – cf. Table 6.

XI.2. The horizontal dimension of the RDA

The horizontal dimension has seven columns separated from the first two working indices by an emphasised vertical line. One row in columns 3.–7. presents itself DA a single keyword article. Let us take a closer look at it.

Column 3. Index a fronte: original lemma with quotation mark

This is the primary and central vertical of the index with the original graphic notation modified to lemma form. Words in the index are in bold, but are not linked.

Table 6. Column 3. in RDA

3		
Index a fronte: lemma without quotation mark	Index a fronte: lemma without quotation mark	Index a fronte: lemma with quotation mark
(...)	(...)	(...)
лонгплей	лонглист	лонглист
лонг-плей	лонг-лист	"лонг-лист"
лонг-листер	лонг-листер	лонг-листер
лонглист	лонгплей	лонгплей
лонг-лист	лонг-плей	лонг-плей
(...)	(...)	(...)

When you click on words in columns 1. or 2. a ‘footprint’ remains in the form of a changed font colour with an underline. This facilitates individual use of the resource.

Column 4. Context with word form

Contextual data includes a sentence or several sentences in their original spelling, with the original word form and the graphic form of the DA highlighted in bold – cf. the table below:

Table 7. Column 4. containing the context of use of DA

4
Context
Уже никто из военкоматчиков не пытается воздействовать на призывника высокими фразами о патриотизме, о священном долге перед Родиной и т.д.

Column 5. Source

In a given level, there is the title of the regional newspaper from which the context was taken for the DA. Most titles also have an original volume number assigned.

Column 6. Year

The year is the date of publication of the newspaper issue containing the context of DA use.

Column 7. RFS

The column in question shows the name of the Russian Federation sub-region from where the regional newspaper originates.

Column 8. City

The title of the column indicates the city of publication of the regional newspaper volume indicated in column 7.

Columns 5., 6., 7., 8. are additionally presented in Table 8.:

Table 8. Columns 5., 6., 7., 8. showing metadata: source (title of newspaper), year, RFS and name of the city in this RFS

5	6	7	8
Source	Year	RFS	City
<Пермские новости №31 (1203)	2003	Пермский край	Пермь

Column 9. MAP

The last column has a map of the Russian Federation. The map is hidden. It appears to readers when they point the cursor at the map record. The map in PNG format shows the RFS (in yellow), the city in the RFS (as a point – black square) and the coordinates in the decimal system (latitude and longitude), e.g. 56.6295264 40.5700972. The map has been created using the DATAWRAPPER (datawrapper.de) design technology – cf. the infographic below:

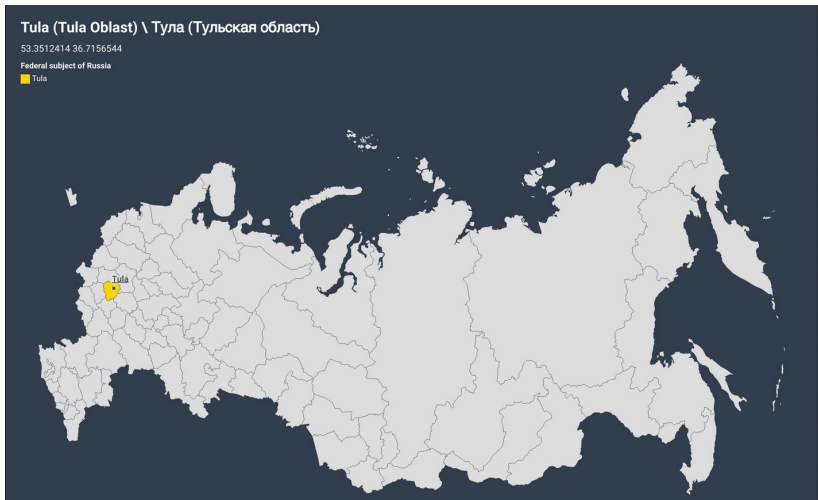


Figure 14. Map of the city of Tula in the Tula region [own study]

You are invited to use the Repository of Dictionary Agnonyms, in its first edition. Please send all comments on the RDA data to jerfed@amu.edu.pl. In addition, I encourage Readers to share ideas related to the expansion and possible optimisation of the RDA.

Bibliography

In this chapter, the notation includes items in several different languages, and in several alphabets. Consequently, the decision has been made to include items in the *Literature* in the Latin alphabet with – in square brackets – the authors' names and titles in the original alphabet. Only the original entries of the items have been included in the *Verification Dictionaries* section.

Literature

- Alatortseva, S.I.** [Алаторцева, С.И.]. **1999.** *Problemy neologii i russkaya neografiya. (Avtoreferat doktorskoy dissertatsii).* [Проблемы неологии и русская неография (Автореферат докторской диссертации)]. Sankt-Petersburg: LGOU im. A.S. Pushkina.
- Aleksandrzak, M. ed.** **2018.** *Języki specjalistyczne w lingwistyce stosowanej: między teorią i praktyką. Seria wydawnicza „Język – Kultura – Komunikacja”, Vol. 23.* Poznań: Wydawnictwo Naukowe UAM.
- Alekseyev, V.A.** [Алексеев, В.А.]. **1973.** *Voprosy stilistiki. Problemy zhurnalistiki 2 (11), 41–49.*
- Babenko, L.G. ed.** [Бабенко, Л.Г.]. **2005.** *Bol'shoy tolkovyy slovar' russkikh sushchestvitel'nykh. Ideograficheskoye opisaniye. Sinonims, antonyms.* [Большой толковый словарь русских существительных. Идеографическое описание. Синонимы, антонимы]. Moskva: AST-PRESS KNIGA.
- Bańko, M., Hebal-Jeziarska M.** **2015.** *Graphic variation in the process of borrowing adaptation. Bulletin de la Société Polonaise de Linguistique XXI, 141–152.*
- Bartkov, V.I., Minina, L.I.** [Бартков, В.И., Минина, Л.И.]. **2019.** *Derivatsionnyy (kolichestvennyy "slovar'" affiksov) sovremennogo russkogo yazyka: v 2 t. T. 1: Trista prepozitivnykh derivatsionnykh morfem]. [Дериватарий (количественный "словарь" аффиксов) современного русского языка: в 2 т. Т. 1: Триста препозитивных деривационных морфем]. Kazan: Buk.*
- Bartmiński, J.** **1980.** *Słownik ludowych stereotypów językowych. Zeszyt próbny.* Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego.
- Bartwicka, H., Fedorushkov, Y., Malek, E., Wawrzyńczyk, J.** **2007.** *Slovarnyye agnonimy russkogo yazyka.* [Словарные агнонимы русского

- языка]. *Series: Semiosis Lexicographica* XLIII. Warsaw: Takt Publishing House.
- Bieleń, S. ed. 1997.** *Looking to the East from the Problems of the Eastern Policy of the Third Republic of Poland.* Warsaw: Centre for Eastern Studies of the University of Warsaw.
- Bierwiazzonek, B. 2006.** *O języku ucieleśnionym.* In: O. Sokolowska, D. Stanilewicz, ed. *Cognitive linguistics III. Językoznawstwo kognitywne III. Kognitywizm w świetle innych teorii.* Gdańsk: Wydawnictwo Uniwersytetu Gdańskiego, 444–479.
- Bortnowska, P. [Бортновска, П.]. 2016.** Defisnye substantivno-substantivnye konstrukcii lichnoj semantiki v sovremennom russkom jazyke. [*Дефисные субстантивно-субстантивные конструкции личной семантики в современном русском языке*]. *Dysertacje Wydziału Neofilologii UAM w Poznaniu* [monograph series], Vol. 17, <https://doi.org/10.14746/9788394719869> – WEB: 07.02.2022.
- Czubiński, A. 1997.** *Europa XX wieku.* Poznań: Wydawnictwo Poznańskie.
- Donaj, Ł. 2001.** *Media w systemie politycznym w Federacji Rosyjskiej.* Poznań: Wydawnictwo UAM.
- Drawicz, A. 1997.** Przewartościowanie w Rosji w polityce III RP. In: S. Bieleń, ed. *Patrząc na Wschód z problematyki polityki wschodniej III RP.* Warszawa: Centrum Badań Wschodnich Uniwersytetu Warszawskiego, 48–62.
- Fedorushkov, Y. 2005.** Automatyczna analiza morfologiczna jako narzędzie oceny poprawności wyrazów języka rosyjskiego. *Investigationes Linguisticae*, Vol. XII, 33–43.
- Fedorushkov, Y. 2007.** *Methods for electronic excerption of new words in Russian.* In: B. Lewandowska-Tomaszczyk, ed. *PALC 2007: Practical Applications in Language and Computers. Papers from the International Conference at the University of Łódź, 19–22 April 2007.* Frankfurt am Main: Peter Lang Verlag, 163–196.
- Fedorushkov, Y. 2008.** *Metody ekscerpji neologizmów języka rosyjskiego (na materiale współczesnej prasy rosyjskiej).* [Doctoral thesis; supervisor – P. Wierzchoń], <http://hdl.handle.net/10593/16715> – WEB: 11.05.2023.
- Fedorushkov, Y. 2009.** Ekscerpja wyrazów cudzojęzycznych z tekstów prasy rosyjskiej. *Series: Semiosis Lexicographica*, Vol. LII, Warszawa: Wydawnictwo Takt..
- Fedorushkov, Y. 2009.** *Słowniki elektroniczne oraz środowisko hipertekstowe w leksykografii a Słownik języka polskiego pod red. W. Doroszewskiego.* In: J. Wawrzyńczyk, ed. *Czterdzieści lat minęło nad Słownikiem Doroszewskiego*, Warszawa: Wydawnictwo Takt, 27–37.

- Fedorushkov, Y. 2010.** Metody automatyzacji ekscerpcji konstrukcji atrybutywnych języka rosyjskiego. In: W. Chlebda, ed. *Na tropach reproduktów. W poszukiwaniu wielowrazowych jednostek języka*, Opole: Wydawnictwo Uniwersytetu Opolskiego, 59–85.
- Fedorushkov, Y. 2019.** *Adaptacja warsztatu leksykograficznego do automatyzacji rozpoznawania frazematów w systemie uczącym się. Na przykładzie konstrukcji werbo-nominalnych języka rosyjskiego*. Poznań: Wydawnictwo Naukowe UAM.
- Fedorushkov, Y. 2018a.** *Kieszonkowy słownik piłki nożnej rosyjsko-polski i polsko-rosyjski*. Warszawa: Centrum Polsko-Rosyjskiego Dialogu i Porozumienia.
- Fedorushkov, Y., 2018b.** Prolegomena do tagowania frazematów w równoległym korpusie rosyjsko-polskim (literatura piękna) w aspekcie przekładoznawczym. *Acta Polono-Ruthenica*, Vol. 2/XXIII, 55–74.
- Fedorushkov, Y., 2018c.** Zwroty czasownikowo-rzeczownikowe o tematyce futbolowej w słowniku rosyjsko-polskim /polsko-rosyjskim. *Studia Rossica Posnaniensia*, Vol. XLIII, 45–62.
- Fedorushkov, Y., Dzienisiewicz, D. 2014.** Automatyzacja wizualizacji grafowej synonimów dla potrzeb dydaktyki języków obcych (na przykładzie przymiotników rosyjskich z prefiksem без-/без-). *Kultury Wschodniosłowiańskie – Oblicza i Dialog*, Vol. 4. 35–48.
- Fedorushkov, Y., Dzienisiewicz, D. 2019.** Metoda wyodrębniania odtwarzalnych konstrukcji werbo-nominalnych języka rosyjskiego z korpusów tekstowych. *Studia Rossica Posnaniensia*, Vol. 44(2), 379–394.
- Fedorushkov, Y., Galinskiy, F., Yurkevich, D. 2018.** *Prolegomena do automatyzacji rozpoznawania rosyjskich frazematów werbo-nominalnych*. In: M. Aleksandrak, ed. 2018. *Języki specjalistyczne w lingwistyce stosowanej: między teorią i praktyką*. Seria wydawnicza „Język – Kultura – Komunikacja”, Vol. 23. Poznań: Wydawnictwo Naukowe UAM, 93–129.
- Fedorushkov, Y., Narloch, A. 2014.** Prolegomena do dydaktycznej prezentacji konceptu językowego w wizualizacji grafowej (na przykładzie rosyjskiego konceptu BELYJ), *Studia Rossica Gedanensia*, Vol. 1, 179–208.
- Fedorushkov, Y., Shutkovsky, T. [Федорущков, Ю., Шутковский, Т.]. 2010.** Leksiko-grammaticheskaja sochetajemost' atributivnyh slovosochetanj russkogo jazyka v kontekste metodov komp'juternoj jekscerpicii. [Лексико-грамматическая сочетаемость атрибутивных словосочетаний русского языка в контексте методов компьютерной экскерпции.]. In: T.R. Guzman, L. Sokolova, I. Voytakova, eds., II. *Mezhdunarodnaya konferentsiya „Russkij jazyk i literatura v mezhdunarodnom obrazovatel'nom prostranstve: sovremennoe sostojanie i perspektivy”*, Gra-

nada, 8–10 sentyabrya 2010 g., Vol. II. Granada: Rubiños – 1860, S.A, 1564–1569.

- Gabdreeva, N.V., Gurčiani, M.T.** [Габдреева, Н.В., Гурчиани, М.Т.]. **2012.** *Slovar' kompozitov ruskogo yazyka novejšego perioda.* [словарь композитов русского языка новейшего периода.]. Moskva: FLINTA / Nauka.
- Handke, K.** **1976.** Budowa morfologiczna i funkcje compositów polskich (z uwzględnieniem innych języków zachodniosłowiańskich). *Prace Językoznawcze*, Vol. 81. Wrocław – Zakład Narodowy im. Ossolińskich. *Prace Językoznawcze*, Vol. 81. Wrocław: Zakład Narodowy im. Ossolińskich.
- Hluszko, A.** **2016.** Wstrząsająca treść jako manipulacyjny element dyskursu konfliktu. *Oblicza komunikacji*, Vol. 9, 135–143.
- Hrynkiewicz-Admaskich, B.** **2005.** *Rosyjsko-polski słownik innowacji leksykalnych.* Poznań: Wydawnictwo Naukowe UAM.
- Jodłowski, S.** **1936.** Z zagadnień związanych z nauczaniem gramatyki polskiej. Wyrazy z. przedrostkami oraz. tzw. zrosty i złożenia. *Język Polski*, Vol. XVI, 103–110.
- Kaliszan, J.** [Калишан, Е.]. **1980.** *Internatsional'nyye prepozitivnyye morfe-mygreko-latinskogo proiskhozhdeniya v sovremennom ruskom i pol'skom slovoobrazovanii.* [Интернациональные препозитивные морфемы греко-латинского происхождения в современном русском и польском словообразовании]. Poznań: Wydawnictwo Naukowe UAM.
- Kobozeva, I.M.** [Кобозева, И.М.]. **2000.** *Leksicheskaya Semantika.* [Лексическая семантика.]. Moskva: Editorial URSS.
- Kostomarov, V.G.** [Костомаров, В.Г.]. **1971.** *Russkiy yazyk na gazetnoy polose. Nekotoryye osobennosti yazyka sovremennoy gazetnoy publiksistiki.* [Русский язык на газетной полосе. Некоторые особенности языка современной газетной публицистики]. Moskva: Izdatel'stvo Moskovskogo universiteta.
- Kostomarov, V.G.** [Костомаров, В.Г.]. **1998.** *Âzykovej vkus èpohi.* [Языковой вкус эпохи.]. *Russkiy yazyk v shkole*, Vol. 1, 11–13.
- Kostomarov, V.G.** [Костомаров, В.Г.]. **1999.** *Âzykovej vkus èpohi. Iz nablyudeniy nad rechevoy praktikoy mass-media.* [Языковой вкус эпохи. Из наблюдений над речевой практикой масс-медиа.]. Sankt-Peterburg: Zlatoust.
- Kotelova, N.Z.** [Котелова, Н.З.]. **2015.** *Izbrannye raboty.* [Избранные работы]. Sankt-Peterburg: Izdatel'stvo Nestor-Istoruya.
- Kotelova, N.Z.** [Котелова, Н.З.]. **ed. 1995.** *Slovar' novyh slov ruskogo âzyka (seredina 50-h – seredina 80-h godov).* [Словарь новых слов русского языка (середина 50-х – середина 80-х годов.]. Sankt-Peterburg: Dmitry Bulanin.

- Kozhina, M.N.** [Кожина, М.Н.]. 2003. Rechevaya sistemnost' funktsional'nogo stilya. In: Kozhina, M.N. ed., 2003, 5–20.
- Kozhina, M.N.** [Кожина, М.Н.]. ed. 2003. *Stilisticheskiy-entsiklopedicheskiy slovar' sovremennogo russkogo yazyka*. Moskva: Flinta, Nauka.
- Krasil'shik, I.S., Rahilina, E.V.** [Красильщик, С.И., Рахилина, Е.В.]. 1992. Predmetnye imena v sisteme „Leksikograf". [Предметные имена с системе "Лексикограф"]. *NTI*, vol. 9, iss. 2, 24–31.
- Kustova, G., Lyashevskaya, O., Paducheva, Ye., Rakhilina, Ye., Shemanaeva, O.** [Кустова, Г., Ляшевская, О. Падучева, Е., Рахилина, Е., Шеманаева, О.]. 2005. *Semanticheskaya razemtka leksiki v Natsional'nom korpusе russkogo yazyka: printsipy, problemy, perspektivy*. [Семантическая раземтка лексики в Национальном корпусе русского языка: принципы, проблемы, перспективы.]. In: V.A. Plungyan, ed. 2005. Национальный корпусского языка: 2003–2005. Результаты и перспективы. Moskva: Indrik.
- Kustova, G., I. Paducheva, Ye.V.** [Кустова, Г.И., Падучева, Е.В.]. 1994. Slovar' kak leksičeskaâ baza dannyh. [Словарь как лексическая база данных.]. *Voprosy âzykoznanîâ*, Vol. 4, 96–106.
- Malcówna, M.** 1974. Dwuczłonowe przezwiska i przydomki staropolskiej. *Onomastica*, Vol. XIX, 191–227.
- Mandrikova, G.M., Krotova, A.G.** [Мандрикова, Г.М., Кротова, А.Г.]. 2013. *Sovremennyy russkiy yazyk: leksikologiya vskhemakh, zadaniyakh i uprazhneniyakh: uchebno-metodicheskoye posobiye*. [Современный русский язык: лексикология в схемах, заданиях и упражнениях: учебно-методическое пособие.]. Novosibirsk: Izdatel'stvo NGTU.
- Marcińczuk, M., Kocoń, J. Ptak, M., Kaczmarek, A.** 2010. *Inforex*, CLARIN-PL digital repository, <http://hdl.handle.net/11321/13> – WEB: 12.12.2022 (web resource – CLARIN-PL digital repository).
- Marcińczuk, M., Oleksy, M., Kocoń, J.** 2017. *Inforex – a Collaborative System for Text Corpora Annotation and Analysis*. In: R. Mitkov, G. Angelova, eds. *Proceedings of the International Conference Recent Advances in Natural Language Processing, RANLP 2017*. Varna: INCOMA Ltd, 473–482.
- Maziarz, M., Piasecki, M., Szpakowicz, S., Rabięga-Wiśniewska, J.** 2011. Semantic Relations Among Nouns in Polish Wordnet Grounded in Lexicographic and Semantic Tradition. *Cognitive Studies/Études Cognitives*, Vol. 11, 161–181.
- Mietla, J.** 2001. Zjawisko hybrydyzacji w czeskiej literaturze językoznawczej. *Bohemistyka*, Vol. 1, 52–60.
- Misturska-Bojanowska, J.** [Мистурса-Бояновска, Й.]. 2013. *Dinamiche-skiye protsessy v slovoobrazovanii imen sushchestvitel'nykh v russkom*

- yazyke 2-oy pol. XX veka*. [Динамические процессы в словообразовании имен существительных в русском языке 2-ой пол. XX века]. Szczecin: Wydawnictwo Naukowe Uniwersytetu Szczecińskiego.
- Morkovkin, V.V., Morkovkina, A.V.** [Морковкин, В.В., Морковкина, А.В.]. **1997.** *Russkie agnonimy: Slova, kotorye my ne znaem*. [Русские агнонимы. Слова, которые мы не знаем]. Moskva: Russkiy yazyk.
- Nemchenko, V.N.** [Немченко, В.Н.]. **1984.** *Sovremennyy russkiy yazyk. Slovoizmeneniye*. [Современный русский язык. *Slovoobrazovaniye*]. Moskva: Vysshaya shkola.
- Novikova, O.V.** **2002** [Новикова, О.В.]. *Strukturno-funksional'naya kharakteristika russkikh neologizmov dvukh stilisticheskikh podsystem kontsa XX – nachala XXI vv. (na materiale gazetnogo diskursa i komp'yuternogo slenga)*. [Структурно-функциональная характеристика русских неологизмов двух стилистических подсистем конца XX – начала XXI вв. (на материале газетного дискурса и компьютерного сленга)]. [thesis, supervisor – O. Lagut.] Novosibirsk: Novosibirskij Gosudarstvennyj universitet, http://zipsites.ru/books/rus_neologizmy – WEB: 13.12.2005.
- Panov, M.V.** [Панов, М.В.]. **1988.** *Iz nablyudeniy nad stilem segodnyashney periodiki. Yazyk sovremennoy publitsistiki*. [Из наблюдений над стилем сегодняшней периодики. Язык современной публицистики]. Moskva: Nauka.
- Plungyan, V.A.** [Плунгян, В.А.]. **ed. 2005.** *Natsional'nyy korpus russkogo yazyka: 2003–2005. Rezul'taty i perspektivy*. [Национальный корпус русского языка: 2003–2005. Результаты и перспективы.]. Moskva: Indrik.
- Plungyan, V.A.** [Плунгян, В.А.]. **ed. 2009.** *Natsional'nyy korpus russkogo yazyka: 2006–2008. Rezul'taty i perspektivy*. [Национальный корпус русского языка: 2006–2008. Новые результаты и перспективы.]. Sankt-Petersburg: Nestor-Istoriya.
- Popov, R.N.** [Попов., Р.Н.]. **1996.** *Novyye slova i slovosochetaniya v yazyke sovremennoy pressy*. [Новые слова и словосочетания в языке современной прессы.]. *Russkiy yazyk v shkole*, Vol. 1, 15–16.
- Popova, T.V.** [Попова., Р.Н.]. **2005.** *Russkaya neologiya i neografiya*. [Русская неология и неография]. Ekaterinburg: UGTU-UI, http://window.edu.ru/window_catalog/files/r28514/ustu121.pdf – WEB: 13 06 2005.
- Popova, T.V., Raciburskaâ L.V., Gugunava, D.V.** [Попова, Т.В., Рацибурская Л.В., Гугунава, Д.В.]. **2005.** *Neologiya i neografiya sovremennogo russkogo yazyka. Uchebnoye posobiye*. [Неология и неография современного русского языка. Учебное пособие.]. Moskva: Flinta, Nauka.
- Pospiszyl, A.** **2000.** *Composita rzeczownikowe w słowniku gwarowym Andrzeja Cinciały*. *Prace Językoznawcze*, Vol. 2, 139–148.

- Puzynina, J. 1968** Co to jest onomazjologia? *Lingua Poznaniensis*, Vol. XII/ XIII, 119–124.
- Rakhilina, Ye.V., Kustova, G.I., Lyashevskaya, O.N., Reznikova, T.I., Shemanayeva, O.Yu.** [Рахилина, Е.В., Кустова, Г.И., Ляшевская, О.Н., Резникова, Т.И., Шеманаева, О.Ю.]. **2009.** *zadachi i principy semantičeskoj razmetki leksiki v NKRÂ.* [Задачи и принципы разметки лексики в НКРЯ.]. In: V. A. Plungân, ed. *Nacionalnyj korpus ruskogo âzyka: 2006–2008. Novye rezul'taty i perspektivy.* Sankt-Petersburg: Nestor, Istorija., 215–239.
- Rahmanova, L.I., Suzdal'tseva, V.N.** [Рахманова, Л.И., Суздальцева, В.Н.]. **1997.** *sovremennyy russkiyyazyk. Leksika. Frazeologiya. Morfoloģiya: Uchebnoye posobiye.* [Современный русский язык. Лексика. Фразеология. Морфология: Учебное пособие.]. Moskva: Izdatel'stvo Moskovskogo Gosudarstvennogo Universiteta.
- Rozental', D.Ë.** [Розенталь, Д.Е.]. **1977.** *Praktičeskaâ stilistika ruskogo âzyka.* [Практическая стилистика русского языка.]. Moskva: Vysshaya shkola.
- Rudnik-Karwatowa, Z. 2003.** O najnowszych tendencjach w słowotwórstwie słowiańskim (na przykładzie języka polskiego i rosyjskiego), *Prace Slavistyczne (Procesy innowacyjne w językach słowiańskich)*, Vol. 114, 199–210.
- Sinyachkina, L.** [Синячкина, Л.]. **1997.** *Osnovnye tendencii v razvitii leksiki âzyka gazety na sovremennom ètape (1985–1996 gg.). Avtoreferat dissertacii.* [Основные тенденции в развитии лексики языка газеты на современном этапе, 1985–1996 гг. Автореферат диссертации...]. Moskva: Moskovskij Gosudarstvennyj Universitet.
- Sklârevskaâ, G.N.** [Сляревская, Г.Н.]. **ed. 2005.** *Tolkovyj slovar' ruskogo âzyka načala XXI veka. Aktual'naâ leksika: ok. 8500 slov i ustojčivyh slovosocetanij.* [Толковый словарь русского языка начала XXI века. Актуальная лексика: около 8500 слов и устойчивых словосочетаний]. Sankt-Peterburg: Èksmo.
- Sosnowski, W. 2011.** *Tendencje analityczne we współczesnym języku polskim i rosyjskim.* [PhD thesis; supervisor – Violetta Koseska-Toszewa]. Warsaw: POLSKA Akademia Nauk/ Instytut Slavistyki – WEB: <https://hdl.handle.net/20.500.12528/42> – WEB: 21 02 2023.
- Trávníček, F. 1948.** *Mluvnice spisovné češtiny*, T I. Praha: Slovanské nakladatelství.
- Trofimova, G.N.** [Трофимова, Г.Н.]. **2001a.** Как обозначит' принадлежность к Internetu. [Как обозначить принадлежность к Интернету.]. *Mir ruskogo slova*, Vol. 4, 73–75.
- Trofimova, G.N.** [Трофимова, Г.Н.]. **2001b.** *K voprosu o spetsifike funkcionirovaniya ruskogo yazyka v Internete (norma i uzus).* [К вопросу

- о специфике функционирования русского языка в Интернете (норма и узус)]. In: *Trudy Mezhdunarodnogo seminaru Dialog-2001 po komp'yuternym tekhnologiyam*. <http://www.dialog-21.ru> – WEB: 06.12.2009.
- Veprava, I.T.** [Вепрева, И.Т.]. **2005.** *Yazykovaya refleksiya v post-sovetskuyu epokhu*. [Языковая рефлексия в постсоветскую эпоху.]. Moskva: OLMA-PRESS.
- Waszakowa, K.** **2005.** *Przejawy internacjonalizacji w słowotwórstwie współczesnej polszczyzny*. Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.
- Waszakowa, K.** **2015.** Kognitywistyczne ujęcia derywatów słowotwórczych a koncepcja Miloša Dokulila. *Biuletyn Polskiego Towarzystwa Językoznawczego*, Vol. LXXI, 111–126.
- Waszakowa, K.** **2021.** *Język w działaniu i działania na języku: szkice semantyczno-słowotwórcze*. 1st ed. Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.
- Wawrzyńczyk, J., Fedorushkov, Y., Ignasiak, A., Kulpina, V., Rodak, A., Siemanowska, U., Tatarinov, V.** *Nazvaniya zhenshchin. Dopolnitel'nyy slovar'*. [Названия женщин. Дополнительный словарь.]. Series: *Semiosis Lexicographica*, Vol. XXXV. Warsaw: Takt Publishing House, 2006.
- Wierzchoń, P.** **2008.** *Fotodokumentacja. Chronologizacja. Emendacja. Teoria i praktyka weryfikacji materiału leksykalnego w badaniach lingwistycznych*. Poznań: SORUS.
- Wierzchoń, P.** **2009.** *Fotodokumentacja 3.0*. In: P. Nowak, P. Nowakowski, eds. *Język. Komunikacja. Informacja*, Vol. 4, 63–80.
- Wierzchoń, P.** **2010.** *Torując drogę teorii lingwochronologizacji, Investigationes Linguisticae*, Vol. XX, 105–185.
- Wierzchoń, P., ed.,** **2010.** *Depozytorium leksykalne języka polskiego. Nowe fotomateriały z lat 1901–2010 – Tom 1*. Warsaw: BEL Studio.
- Wierzchoń, P., Graliński F.** **2016.** Z kart historii „parcia na” neologizmy, *Poradnik Językowy*, Vol. 4, 110–129.
- Zaliznyak, A.A.** [Зализняк, А.А.]. **1987.** *Grammaticheskii slovar' russkogo yazyka. Slovoizmeneniye: okolo 100000 slov*. [Грамматический словарь русского языка. Словоизменение: около 100 000 слов.]. Moskva: Russkiy yazyk.
- Zaliznyak, A.A.** [Зализняк, А.А.]. **2022.** *Grammaticheskii slovar' russkogo yazyka. Slovoizmeneniye: okolo 110000 slov*. [Грамматический словарь русского языка. Словоизменение: около 110 000 слов.]. VIth ed. with online version. Moskva: AST-Press, <https://gramdict.ru/> – WEB: 06.12.2022.

Verification dictionaries

БАС (большой академический словарь):

Чернышёв, В.И. et al. 1950–1965. *Словарь современного русского литературного языка*. Т. 1 (А – Б) – VOL. 15 (X – Я). Москва / Ленинград: Издательство Академии Наук СССР.

БАС – continued:

Чернышёв, В.И. 1951. *Словарь современного русского литературного языка* (В).

Чернышёв, В.И. 1954. *Словарь современного русского литературного языка* (Г – Е).

Чернышёв, В.И. 1955. *Словарь современного русского литературного языка* (Ж – З).

Чернышёв, В.И. 1956. *Словарь современного русского литературного языка* (И – К).

Чернышёв, В.И. 1957. *Словарь современного русского литературного языка* (Л – М).

Чернышёв, В.И. 1958. *Словарь современного русского литературного языка* (Н).

Чернышёв, В.И. 1959. *Словарь современного русского литературного языка* (О).

Чернышёв, В.И. 1959. *Словарь современного русского литературного языка* (П – пнуть).

Чернышёв, В.И. 1960. *Словарь современного русского литературного языка* (По – поясочек).

Чернышёв, В.И. 1961. *Словарь современного русского литературного языка* (Пра – пятью).

Чернышёв, В.И. 1961. *Словарь современного русского литературного языка* (Р).

Чернышёв, В.И. 1962. *Словарь современного русского литературного языка* (С – сняться).

Чернышёв, В.И. 1963. *Словарь современного русского литературного языка* (Со – сям).

Чернышёв, В.И. 1963. *Словарь современного русского литературного языка* (Т).

Чернышёв, В.И. 1964. *Словарь современного русского литературного языка* (У – Ф).

Ковтун, Л.С., Петушков, В.П. 1965. *Словарь современного русского литературного языка* (X – Я).

- Брагина, А.А. 1997.** Новое в русской лексике. Словарные материалы – 1988. *Филологические науки* 2, 111–115.
- Буцева, Н.В. ed. 2004.** Новое в русской лексике. Словарные материалы – 1990. Санкт-Петербург: Дмитрий Буланин.
- Плотицын, В., Соловьев, Н., Юдина, А., Судоплатова, М., Бояркина, В., Козулина, Н. 2001.** *Новое в русской лексике. Словарные материалы – 1989.* Санкт-Петербург: Дмитрий Буланин.
- Денисенко, Ю.Ф. ed. 2005.** *Новое в русской лексике. Словарные материалы – 1991.* Санкт-Петербург: Дмитрий Буланин.
- Денисенко, Ю.Ф. ed. 2006.** *Новое в русской лексике. Словарные материалы – 1994.* Санкт-Петербург: Дмитрий Буланин.
- Евгеньева, А.П. ed. 2001.** *Словарь синонимов русского языка.* Т. I–II. Москва: Астрель.
- Зализняк, А.А. 1987.** *Грамматический словарь русского языка. Словоизменение: около 100000 слов.* 3rd ed. modified and supplemented. Москва: Русский язык.
- Зализняк, А.А. 2022.** *Грамматический словарь русского языка. Словоизменение: около 110000 слов.* 6th ed. Москва: АСТ-Пресс. Official online version of the dictionary: <https://gramdict.ru/> – WEB: 06.12.2022.
- Котелова, Н.З. ed. 1980.** *Новое в русской лексике. Словарные материалы – 1977.* Москва: Русский язык.
- Котелова, Н.З. ed. 1981.** *Новое в русской лексике. Словарные материалы – 1978.* Москва: Русский язык.
- Котелова, Н.З. ed. 1982.** *Новое в русской лексике. Словарные материалы – 1979.* Москва: Русский язык.
- Котелова, Н.З. ed. 1984a.** *Новое в русской лексике. Словарные материалы – 1980.* Москва: Русский язык.
- Котелова, Н.З. ed. 1984b.** *Новые слова и значения. Словарь-справочник по материалам прессы и литературы 70-х гг.* Москва: Русский язык.
- Котелова, Н.З. ed. 1986a.** *Новое в русской лексике. Словарные материалы – 1981.* Москва: Русский язык.
- Котелова, Н.З. ed. 1986b.** *Новое в русской лексике. Словарные материалы – 1982.* Москва: Русский язык.
- Котелова, Н.З. ed. 1987.** *Новое в русской лексике. Словарные материалы – 1983.* Москва: Русский язык.
- Котелова, Н.З. ed. 1989.** *Новое в русской лексике. Словарные материалы – 1984.* Москва: Русский язык.
- Котелова, Н.З. ed. 1995.** *Словарь новых слов русского языка (середина 50-х-середина 80-х гг.).* Санкт-Петербург: Дмитрий Буланин.

- Котелова, Н.З. ed. 1996.** *Новое в русской лексике. Словарные материалы – 1986.* Санкт-Петербург: Дмитрий Буланин.
- Котелова, Н.З., Денисенко, Ю.Ф. eds. 1996a.** *Новое в русской лексике. Словарные материалы – 1985.* Москва: Дмитрий Буланин.
- Котелова, Н.З., Денисенко, Ю.Ф. eds. 1996б.** *Новое в русской лексике. Словарные материалы – 1987.* Москва: Дмитрий Буланин.
- Котелова, Н.З., Сорокин, Ю.С. ed. 1971.** *Новые слова и значения. Словарь-справочник по материалам прессы и литературы 60-х гг.* Москва: Советская Энциклопедия.
- Левашов, Е.А. ed. 1996.** *Новое в русской лексике. Словарные материалы – 1988.* Санкт-Петербург: Дмитрий Буланин.
- Левашов, Е.А. ed. 1997.** *Новые слова и значения. Словарь-справочник по материалам прессы и литературы 80-х гг.* Санкт-Петербург: Дмитрий Буланин.
- Левашов, Е.А. ed. 2004.** *Новое в русской лексике. Словарные материалы – 1992.* Санкт-Петербург: Дмитрий Буланин.
- МАС** (малый академический словарь):
- Евгеньева, А.П. ed. 1981.** *Словарь русского языка. В четырёх томах.* Т. I (А – Й). Ed. II. Москва: Русский язык.
- Евгеньева, А.П. ed. 1982.** *Словарь русского языка. В четырёх томах.* Т. II (К – О). Ed. II. Москва: Русский язык.
- Евгеньева, А.П. ed. 1983.** *Словарь русского языка. В четырёх томах.* Т. III (П – Р). Ed. II. Москва: Русский язык.
- Евгеньева, А.П. ed. 1984.** *Словарь русского языка. В четырёх томах.* Т. IV (С – Я). Ed. II. Москва: Русский язык.
- Милейковская, Г., Казем-Бек, Г., Домбровская И., Гаевская, Л. 1988.** *Материалы для словаря русского языка 70-х и 80-х годов XX века не зарегистрированные в нормативных словарях.* Т. 1. Warszawa: Wydawnictwo Uniwersytetu Warszawskiego.
- Милейковская, Г., Петельщик, Р., Тарновская, И., Тимошук, М. 1996.** *Материалы для словаря русского языка 80-х годов XX века не зарегистрированные в нормативных словарях.* Т. 3. Warszawa: Energeia.
- Милейковская, Г., Петельщик, Р., Тимошук, М. 1990.** *Материалы для словаря русского языка 70-х и 80-х годов XX века не зарегистрированные в нормативных словарях.* Т. 2. Warszawa: Wydawnictwo Uniwersytetu Warszawskiego.
- Ожегов, С.И., Шведова Н.Ю. 1992.** *Толковый словарь русского языка.* Москва: Русский язык.
- Ушаков, Д.Н. ed. 1996.** *Толковый словарь русского языка.* Москва: ТЕРРА.

Ушаков, Д.Н., Крючков, С.Е. 1980. *Орфографический словарь*. Москва: Просвещение.

Corpora and other textual resources

Piasecki, M. 2018. *WordnetLoom 2.0* – WEB: <http://hdl.handle.net/11321/557->
WEB: 12.12.2022 (web resource – CLARIN-PL digital repository).

Pawlowski, A.T. 2016. *ChronoPress – Chronologica Corpus*, <http://hdl.handle.net/11321/260> – WEB: 12.12.2022 (web resource – CLARIN-PL digital repository).

Pawlowski, A. 2021. *ChronoPress – Corpus of Press Texts*, doi:10.34616/139101. <http://chronopress.clarin-pl.eu/> – WEB: 12.12.2022 (web resource – CLARIN-PL digital repository).

Fedorushkov, Y. 2023. *Corpus of Russian Local Press of the Millennium Period (1996–2006)*, CLARIN-PL digital repository – WEB: <http://hdl.handle.net/11321/911>. WEB: 22.04.2023:

Polikarpov, A.A. 2007. *Komp'yuternyy korpus tekstov russkikh gazety kontsa XX-ogo veka.* [Компьютерный корпус текстов русских газет конца XX-ого века.]. – WEB: https://www.philol.msu.ru/~lex/main_en.htm – WEB: 23.12.2023.

The National Corpus of the Russian language. 2022. *Semantičeskâ razmetka. O leksiko-semantičeskoj informacii v Korpuse.* [Семантическая разметка. О лексико-семантической информации в Корпусе.]. <https://ruscorpora.ru/page/instruction-semantic/> – WEB: 12.01.2023.

Index of tables

Table 1. Example of key data in a keyword article in RDA.....	16
Table 2. Linguoepochronological data in RDA: RFS and DA.....	22
Table 3. Path to the DA list	44
Table 4. RDA columns in vertical (columns 1., 2.) and horizontal (3.–9.) dimensions.....	79
Table 5. A tergo and a fronte indexes with two-sided hyperlinks – columns 1. and 2.....	82
Table 6. Column 3. in RDA	83
Table 7. Column 4. containing the context of use of DA.....	84
Table 8. Columns 5, 6, 7, 8 showing metadata: source (title of newspaper), year, RFS and name of the city in this RFS	84

Index of illustrations

Figure 1. Explosion in the growth of unique words in the millennium decade (individual regions and years) [own study].....	17
Figure 2. Locations of IDA word occurrences on the map of regions [own study].....	42
Figure 3. Occurrence of DA in the local press in different RFSs in 1996 [own study].....	61
Figure 4. Occurrence of DA in the local press in different RFSs in 1997 [own study].....	62
Figure 5. Occurrence of DA in the local press in different RFSs in 1998 [own study].....	62
Figure 6. Occurrence of DA in the local press in different RFSs in 1999 [own study].....	63
Figure 7. Occurrence of DA in the local press in different RFSs in 2000 [own study].....	63
Figure 8. Occurrence of DA in the local press in different RFSs in 2001 [own study].....	64
Figure 10. Occurrence of DA in the local press in different RFSs in 2003 [own study].....	65
Figure 11. Occurrence of DA in the local press in different RFSs in 1994 [own study].....	65
Figure 12. Occurrence of DA in the local press in different RFSs in 2005 [own study].....	66

Figure 13. Occurrence of DA in the local press in different RFSs in 2006
[own study].....66
Figure 14. Map of the city of Tula in the Tula region [own study].....85

REPOZYTORIUM KONTEKSTOWE
AGNONIMÓW SŁOWNIKOWYCH
W ROSYJSKIEJ PRASIE LOKALNEJ
(1996–2006)

10,000 HASEŁ W FORMACIE HTML

Pragnę podziękować doktorowi inżynierowi Mateuszowi Lango (Politechnika Poznańska, Wydział Informatyki i Telekomunikacji, Instytut Informatyki) za pomoc w rozwiązaniu problemów programistycznych, doktorowi Maciejowi Waraczewskiemu (dyrektorowi Centrum Językowego Poligłoci w Poznaniu) za pomoc w kontekstualizacji agnonimów, zespołowi CLARIN za pomoc w publikacji korpusu gazet – szczególnie magistrowi inżynierowi Tomaszowi Naskrętowi i doktorowi Janowi Wieczorkowi, informatykom Maciejowi Danielakowi oraz Michałowi Rachwalskiemu z Wydziału Neofilologii UAM za wskazówki dotyczące kodu skryptów PYTHON oraz za pomoc w obróbce graficznej pliku Repozytorium w formacie HTML.

Przedmowa

Znany fakt wskazywany przez badaczy stanowi to, że jedną z aktualnych cech leksykalnych języków słowiańskich jest ich internacjonalizacja (Mietła 2001: 52, Kaliszan 1980, Waszakowa 2005: 10, 26, 28, 206). W literaturze przedmiotu mówi się przede wszystkim o otwartości języka rosyjskiego, podobnie jak innych języków słowiańskich (Waszakowa 2005: 194; por. tytuł rozdziału: *Tendencja do internacjonalizacji jako zjawisko ogólnosłowiańskie*) na inne języki (Trofimova 2001a, Alatorseva 1999). Badaczka S.I. Alatorseva uważa, że neologizmy to przede wszystkim przejaw wewnętrznych czynników kulturowych: „(...) Proces neologizacji jest nieodłączny w mniejszym lub większym stopniu we wszystkich językach narodowych, ale każda kultura narodowa ma do niego swój własny, szczególny stosunek. Rosja, naszym zdaniem, jest jednym z krajów, w których stosunek do innowacji językowych jest zbyt demokratyczny. Głównym problemem tego kierunku jest kultura narodowa i neologizmy obcojęzyczne.” (Alatorseva 1999: 9). Leksyka bogaci się i modyfikuje poprzez proces zapożyczania. Pojawia się leksyka nowa, nienotowana wcześniej w tekstach. Nowe leksemy są właśnie tymi papierkami lakmusowymi, które pozwalają uwydatnić przebieg procesów językowych. Istnieje więc wyraźna potrzeba na stworzenie neograficznego zasobu leksykalnego, umożliwiającego zbadanie zarówno procesu internacjonalizacji języka rosyjskiego w związku ze wspomnianym „otwarcie” Rosji na Zachód w okresie zmiany ustroju politycznego w latach 1996–2006, jak i naukowo uzasadnionych czynności badawczych skierowanych na analizę stricte lingwistyczną. Na obecnym etapie ta otwartość Federacji Rosyjskiej została niewątpliwie zachwiana w związku z odizolowaniem się Rosji od państw Zachodu, zamknięciem granic, wycofaniem firm i korporacji zachodnich, z określeniem państw zachodnich „krajami nieprzyjaznymi”, agresją militarną Rosji w Czeczeni, a później w Gruzji, Syrii, państwach afrykańskich, a także – 24 lutego 2022 roku – w Ukrainie. Duża część periodyków, które figurują w naszym surowcu tekstowym – korpusie gazet rosyjskich – zniknęła lub przyjęła retorykę propagandową. W obecnej Rosji dominują media – w tym prasowe – podporządkowane

językowi i retoryce władzy. Z tego punktu widzenia warto udostępnić zarówno repozytorium nienotowanych wcześniej słów z kontekstami ich użycia, jak i sam korpus gazet, w których ma miejsce dialog, a także zauważalna jest jeszcze wolność myśli i słowa, odrębna od monologu współczesnej maszyny propagandowej. Publiczne udostępnienie takiego zasobu umożliwi rozpoczęcie całej serii prac badawczych związanych z rozwojem leksyki, morfologii, słowotwórstwa oraz semantyki współczesnego języka rosyjskiego.

Repozytorium agnonimów słownikowych (RAS) ma charakter lingwogeochronologiczny. Skierowane jest do szerokiego grona lingwistów – przede wszystkim leksykografów i leksykologów. Źródła słów to archiwa lokalnej prasy rosyjskiej, z których został zbudowany *Korpus Rosyjskiej Prasy Lokalnej Okresu Milenijnego* (KRPLOM). Niniejsze opracowanie prezentuje wynik prac badawczych oraz ekscerpcyjnych wyłącznie w ramach danego korpusu. Należy tu wspomnieć, że dany korpus został upubliczniony: KRPLOM udostępniony jest w systemie DSpace zasobu CLARIN-PL (Fedorushkov 2023).

Przedkładałam autorski wybór 10 tys. artykułów hasłowych ilustrujących przede wszystkim neologizmy rosyjskiej prasy lokalnej z lat 1996–2006 zebranej w poszczególnych regionach Federacji Rosyjskiej. Poszukiwania periodyków, w tym archiwów, odbywało się na zasadzie analogii typu „wersja drukowana = wersja internetowa”, tj. wydrukowanego numeru określonej gazety z odpowiadającą jej wersją online. Z wersji online gazet utworzono korpus tekstów. Komponowanie KRPLOM-u odbyło się w latach 2005–2006. Rok 2006 jest zatem także dolną granicą czasową dla ekscerpcji AS oraz wszelkiej weryfikacji uzyskanych wyrazów.

Wstępna lista (LNS) takich potencjalnych neologizmów pojawiła się w ramach załogarytmizowanych czynności ekscerpcyjnych, które zostały opisane w rozprawie doktorskiej pt. *Metody ekscerpcji neologizmów języka rosyjskiego (na materiale współczesnej prasy rosyjskiej)*¹

¹ Z treścią rozprawy doktorskiej można zapoznać się online w zasobie Repozytorium Uniwersytetu im. Adama Mickiewicza (UAM) w Poznaniu – AMUR: (<https://repozytorium.amu.edu.pl> – WEB: 21.03.2021). Praca została napisana

(Fedorushkov 2008). Manualna ekscerpca nowych słów jest bardzo skomplikowanym procesem. Wiąże się to z prostym faktem: ekscerpca dostępnych ilości tekstowych jest niewykonalna w sposób manualny. Wymagany jest aparat metodyczny umożliwiający automatyzację jej niektórych etapów. W związku z tym niżej pojawi się kilka pojęć technicznych głównie na oznaczenie ewolucji listy LNS do listy AS (por. rozdział IV. Metody badawcze). W celu wydobycia neologizmów należało utworzyć technologię dotarcia do listy wyrazów unikatowych pod względem graficznym – tzw. *neonimów* (neologizmów graficzno-leksykalnych). Ogólna lista słów korpusu to 250 mln okazów. Lista końcowa unikatów graficznych – lista LNS(AK) (por. rozdział IV. Metody badawcze) – liczy 300 tys. ciągów unikatowych – niezlematyzowanych typów. Wybór słów w indeksie repozytorium można określić jako autorski, agnomy były wybierane i weryfikowane w sposób subiektywny. Z listy LNS(AK) manualnie dobrano wyrazy składające się na listę LNS. Jeśli w repozytorium liczba słów z prefiksem *анти-** wynosi kilkadziesiąt, to w jednej z pierwszych list AS wyrazy z *анти-*, *-анти* oraz *-анти-* mogą stanowić osobny słownik – jest ich prawie 2700. Wymaga to jednak osobnych czynności leksykograficznych. Proces ekscerpacji AS można wyrazić za pomocą następującej kolejności:

- KRPLOM (250 mln tokenów);
- Lista LNS(AK) (300 tys. unikatów nieobecnych w słowniku Zaliżnyak 1987 – por. rozdział IV. pt. Metody badawcze sdcscd)
- Lista LNS (lista ponad 15 tys. zlematyzowanych AS bez podania kontekstu występowania, po weryfikacji notowania w odpowiednio dobranych źródłach leksykograficznych). Przykłady słów z Listy LNS:

100-квартирный, 3D-видеоадаптер, АА-батарейка, АВИА-ШОУ, АВИСМА-МетрАТек, Аз-Буки-Веди, АмурЭкспоФорум-2006, о-о-очень, рагга-джангловый, ракетоподобный, рэпообразный, сайтотворчество, свинговать, тарантинофилософия, "телекамикадзе", трендмейкер, ультралайтовый, ультраэпатажный, файлообменный, фильмпроверщица etc.

pod kierunkiem naukowym prof. Piotra Wierzchońa w Instytucie Językoznawstwa UAM (obecnie jest to Wydział Etnolingwistyki). Zaznaczam, że w niniejszej publikacji będą wykorzystane nieliczne fragmenty treści ze wspomnianej dysertacji – zwłaszcza w rozdziale VIII.

- *Indeks IAS* (spis ponad 10 tys. AS poddanych weryfikacji pod względem poprawności w kontekstach użycia oraz dodatkowo ukontekstwowionych, tj. poddanych kontekstualizacji – dodaniu do hasła pojedynczego zdania lub dłuższego fragmentu z kontekstu użycia).

Treść RAS to rezultat ukontekstowania całego indeksu lemm – ponad 10 tys. unikatowych słów z wyróżnieniem odpowiedniego leksemu w kontekście, ze wskazaniem nazwy periodyku, numeru, roku ukazania się, oraz – co najważniejsze: regionu oraz miasta dla każdego escerptu. Powołujemy się tu na opinię eksperta w dziedzinie lingwochronologizacji: Piotr Wierzchoń, jeden ze współautorów dziesięciotomowego *Depozytorium leksykalnego języka polskiego*, zaznacza, że ekscepcja lingwochronologiczna neonimów przewiduje następane niezbędne parametry: hasło, ilustrację tekstową, datowanie, lokalizację geograficzną oraz czasową (por. Wierzchoń, red. 2010).

Mając na uwadze ogrom niuansów technicznych, metodologicznych, merytorycznych związanych z RAS, ograniczam się do umieszczenia szczegółowego opis powstania RAS w następujących rozdziałach. Opiszę przede wszystkim cel szczegółowy utworzenia RAS, ogólne kryteria ekscerpcyjne haseł, etapy pracy nad RAS, kryteria doboru kontekstu oraz procesu kontekstualizacji, kwestię definicji haseł, korpus gazet, okres historyczny, w którym powstały numery periodyków, strukturę RAS, strukturę artykułu hasłowego oraz serię decyzji leksykograficznych w trakcie tworzenia indeksu *a fronte*, a także prolegomenę opisu ogólnych tendencji w rozwoju morfologii wyrazów nowych.

I. Agnonim słownikowy

W swojej książce pt. *Fotodokumentacja. Chronologizacja. Emendacja. Teoria i praktyka weryfikacji materiału leksykalnego w badaniach lingwistycznych* Piotr Wierzchoń tworzy listę zasad (tzw. dyrektyw) chronologizacji słownictwa, przy czym wskazuje na to, że termin neologizm należy używać z wielką ostrożnością w celu uniknięcia fałszywej hipotezy datacyjnej. Jedną z dyrektyw głosi: „W celu minimalizacji liczby fałszywych hipotez datacyjnych należy zrezygnować z używania lingwochronologicznego określenia neologizm w stosunku do jednostek języka uważanych za datowane na okres powojenny na rzecz terminu agnonim słownikowy, który mówi, że danej jednostki języka nie notowano w powojennych słownikach bądź w opracowaniach leksykograficznych” (Wierzchoń 2008: 120). Wobec danego zastrzeżenia zamiast terminów *neologizm*, bądź „bardziej ostrożnego” – *neologizm potencjalny* – lub nawet *neonim* będą stosować termin nadrzędny, mianowicie – *agnonim słownikowy* (AS). Za pomocą danego hiperonimu podkreślam, że danej jednostki języka nie notowano w określonej serii słowników poprzedzających powstanie listy, w której skumulowano ponad 10 tys. AS. Innymi słowy, jest to termin techniczny i strategiczny. W jego zasięgu funkcjonalnym są nie tylko neonimy (neologizmy, okazjonalizmy), ale też klasycznie rozumiane agnonimy (tj. wyrazy rzadkie, mało znane, mało używane, niezrozumiałe – skrajnie peryferyjne), historyzmy oraz hapaks legomena – okazjonalne wyrazy autorskie użyte przez poszczególnych dziennikarzy m.in. w celu intensyfikacji ekspresji. Spośród AS dominują w większej mierze właśnie wyrazy nowe, wygenerowane lub powtórzone na potrzeby współczesności, np.: *SMS-ругательство, смс, онлайн-болтовня, фэшн-неделя, FM-станция, аудиовидеосинхронизация, профайлер, путиноманья, путиномика, брендмейкер, порноюмор*. Neologizm może być zarejestrowany właśnie spośród szeregu agnonimów słowników. Neologizm to jednostka uzualna: uznana może być za nową, jeśli leksykalizuje się i aktywnie funkcjonuje w uzusie, jego obszarach nieperyferyjnych dopiero po „próbie czasu” zgodnie z ideą zawartej w definicji neologizmu autorstwa rosyjskiej badaczki neologii Nadzieždy Kotyolovej.

Są to „słowa, znaczenia słów, idiomy, które istnieją zwykle w pewnym okresie, w pewnym języku, subjęzyku, sferze językowej i nie istnieją w pewnym ograniczonym poprzednim okresie w tym samym języku, subjęzyku, sferze językowej” (Kotyolova 1978: 22). Neologizm wobec tego to wyraz, który „ma zamiar” przestać funkcjonować jako słowo nowe, stając się wyrazem bez suprasemu konotacyjnego ‘nowy’, wyrazem szeregowym, aktywnym. Do tej pory jednak taki sem zawiera, wonec czego uważany jest za novum w leksyce. Powtórzmy: jeśli chcemy uznać wyraz za neologizm, powinniśmy zbadać jego los w kontekstach użycia, frekwencję, zasięg terytorialny, czasowy, dialektyczny, socjolektyczny. Uznanie takiego wyrazu za neologizm zaczyna się od uznania takiego wyrazu za agnonim słownikowy. Właśnie ta idea kieruje ku sformułowaniu celu utworzenia repozytorium.

II. Cel naukowy utworzenia repozytorium

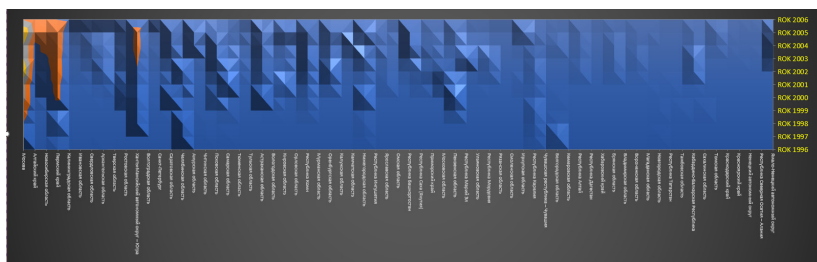
Leksyka w RAS pochodzi z okresu 1996–2006, który technicznie można określić jako dekadę milenijną. Istnieje pogląd (Misturska-Bojanowska 2013, Bortnowska 2016 i in.), że język fleksyjny, jakim jest język rosyjski w badanym okresie rozwoju (koniec XX – początek XXI wieku w Rosji – por. Bieleń 1997, Drawicz 1997, Mandrikova, Krotova 2013) intensywnie kontynuuje dążenie do analityzmu, aglutynizacji, przy czym tendencja ta dotyczy słownictwa z określonego spektrum tematycznego (Handke 1976, Malcówna 1974, Pospiszyl 2000). Dotyczy to przede wszystkim kompozytów (Jodłowski 1936, Bukchina, Kalakuckaya 1974, Waszakowa 2015, Waszakowa 2021, szczególnie dla rosyjskiego Gabdreewa, Gurchiani 2012, Sinyachkina 1997, Vepreva 2005). Tendencje analityzacji i paradygmatyzacji we współczesnej morfotaktyce języka rosyjskiego stanowią etap wzbogacania się określonych sfer leksyki i przyczyniają się do zmiany kierunku rozwoju języka. Wzbogacanie się języka nowym słownictwem – neonomami – jest właśnie tym obszarem, który sygnalizuje te zmiany. Aby umożliwić wgląd do wskazanej tendencji wzrostu neonomów o określonych parametrach, w roku 2009 rozpocząłem komponowanie oraz ukontekstowanie IAS – indeksu agnominów słownikowych (AS) oraz ich wariacji graficznych (por. Bańko, Hebal-Jeziarska 2015) w procesie adaptacji. W roku 2022 ukończyłem czynności weryfikacji datacji oraz przeprowadziłem dobór kontekstu dla każdego wyrazu z indeksu IAS. Dodatkowo każdy AS oprócz kontekstu został opatrzony metadanymi: tytuł gazety lokalnej, datacja (rok), nazwa regionu występowania AS oraz miasto z podaniem mapy geograficznej. Zebrane zostały parametry dla każdego z dziesięciu tysięcy słów w IAS. Na przykład dla hasła "*земфироваться*" są to następujące dane:

Tabela 1. Przykład najważniejszych danych w artykule hasłowym w RAS

TYP DANYCH	ELEMENTY ARTYKUŁ HASŁOWEGO
1. Lematyzowany AS jako hasło wyróżnione w pogrubieniu. Jeśli w kontekście AS figuruje z cudzysłowem, to cudzysłów pozostaje w haśle.	"земфироваться"
2. Kontekst użycia z wyróżnioną formą kontekstową AS	Ю.Т. Положение в русском роке такое же, как и в роке вообще, то есть налицо уход в некие "потусторонние миры", не сказать, чтобы в подполье. Есть волна коммерческого рока, которого все больше, все более мейнстримизуется (мейнстримизируется), "земфируется" , "муммитролляется", "ошнуривается" и так далее.
3. Metadane: tytuł czasopisma, numer i rok wydania, subiekt Federacji Rosyjskiej (RFS), miasto	<Завтра №52 (632);2005; Москва [jako RFS]; Москва [jako miasto]>

Obserwując wyrazy-neonimy z danego zbioru, można zaobserwować lawinową tendencję do (...) afiksacji (prefiksacji / prefiksoidacji / sufiksacji / sufiksoidacji), wzrostu konstrukcji łącznikowych oraz kompozycji (np. złożzeń dwurdzeniowych z co najmniej jednym elementem zapożyczonym, np. *IT-технология, sms-голосование, лонг-дринк, интернет-переписка, смарт-фон, смарттелефон*). Wobec tego wyrazy te mogą być klasyfikowane – w możliwych badaniach – z pozycji najbardziej subtelnej klasyfikacji – tzw. „pentagonicznej” skali: afiks – afiksoid – półafiks – radixoid – radix (rdzeń) (Minina, Bartkov 1980, Bartkov 2006, Bartkov, Minina 2019) z doprecyzowanym algorytmem rozróżnienia poszczególnych formantów pod względem funkcjonalnym.

Badacze wskazują, że język gazet wyklarował się w latach 70.–80. XX wieku (Novikova 2002, Kostomarov 1971, Gogova 1975, Panov 1988), natomiast w okresie milenijnym przeżywa swoje apogeum – kla-ruje się wręcz specyficzny – zliberalizowany, znajdujący się w opozycji do normy literackiej (Kostomarov 1999, Czubiński 1997, Donaj 2001) – socjolekt dziennikarzy gazet lokalnych, który przeżył transfer z gazet drukowanych do wersji cyfrowych, przyczyniając się tym samym do zwiększonego wglądu przez obywateli (Trofimova 2001b). Celem repozytorium RAS jest przede wszystkim wyekspozowanie zarówno samego słownictwa, jak i przyjrzenie się jego ewolucji. Ponadto stwierdzam, że w opisywanym okresie można zaobserwować fenomenalnie wysoki wzrost, wręcz eksplozję przypuszczalnie najnowszych wyrazów niezauważonych w dotychczasowych² słownikach – por. poniższa infografika:



Rysunek 1. Eksplozja wzrostu liczby unikalnych słów w dekadzie milenijnej (poszczególne regiony i lata) [oprac. własne]

Kolejne cele badawcze stawiane uzyskanym danym językowym powinny uwzględniać zatem cechy inherentne tych danych. Mam na myśli oczywiście właściwości morfotaktyczne zarejestrowanych w RAS. Kolejną istotną cechą tych wyrazów jest ich neonimiczna unikatowość. Cecha ta wymaga dogłębnej weryfikacji kwerendowej i konsultacyjnej w poszczególnych leksykonach tekstowych.

² Mam tu myśli wyrazy niezauważone w treści słowników weryfikacyjnych. Ze spisem tym można zapoznać się w podrozdziale pt. *Słowniki weryfikacyjne* w Bibliografii).

Jeszcze raz wyraźnie podkreślę: celem pracy nad RAS nie było określenie m.in. tendencji słowotwórczych współczesnej ruszczyzny, nie skupiałem się również jedynie na przejawach internacjonalizacji w słowotwórstwie języka rosyjskiego badanego okresu. Oczywiście badania „uaktualniające” te zagadnienia możliwe są do przeprowadzenia w dalszej kolejności. Jednakże te badania warunkowane są odpowiednio obszernym materiałem badawczym. Celem pracy nad RAS pracy było dostarczenie możliwie obfitego materiału, a nie jego analiza. Pewne elementy analizy współczesnego języka rosyjskiego naturalnie ujęte są w dysertacji Fedorushkov 2008, w której miał miejsce również opis serii decyzji analitycznych. Dane decyzje dotyczyły kwestii, które z konstrukcji nietypowych można – z pewnymi zastrzeżeniami, np. z parametrem *in status nascendi* – przyjąć jako dopuszczalne w języku rosyjskim, a jakie twory pozostają zdecydowanie błędne; stąd np. wydaje się, że należy uznać za poprawne konstrukcje typu: *девушка*-"*молодогравдейка*", *интернет*-"*ливайс*", *культуртрегер*-"*Фаланстер*", *курсант*-"*вкиушник*", *куртка*-"*бомбер*", *муж*-"*анекдотоман*" itp.). Podsumowując, należy zdecydowanie stwierdzić, że RAS w powyższym sensie ma walor rejestracyjny, choć w praktyce rejestracji nie sposób uchylić się od decyzji o charakterze normatywnym.

Co prezentują sobą metody badawcze oraz ogół czynności z nimi związanych? Zanim odpowiem na to pytanie, w pierwszej kolejności opiszę korpus gazet, z którego został wyekscerpowany materiał słowny. Przejdźmy zatem do omówienia właściwości statystycznych oraz geopolitycznych KRPLOM-u.

III. Korpus KRPLOM i ograniczenia dla RAS

W czasie tworzenia (lata 2005–2006³) KRPLOM-u do Federacji Rosyjskiej (FR) należało 21 republik, 48 regionów, 7 krajów, 1 region autonomiczny, 7 okręgów autonomicznych, 2 miasta o znaczeniu regionalnym (Sankt Petersburg oraz Moskwa): łącznie 86 tzw. *subiektów* (RFS) o znaczeniu federacyjnym⁴.

Uzyskana pełna lista słów KRPLOM na rok 2007 liczyła ponad 5 mln niepowtarzających się słów-typów (ang. *unique words, types*) oraz 250 mln okazów (tokenów). Do korpusu dołączone zostały materiały i archiwa z otwartego publicznego dostępu około 300 gazet z 76 regionów Federacji Rosyjskiej. W pozostałej części regionów nie udało się znaleźć odpowiedników http gazet lokalnych. Materiał KRPLOM stanowią dane prasowe od 1995 roku po rok 2006/pocz. 2007. Dane archiwalne przypadające na wskazane lata nie były symetryczne: archiwum jednej gazety – to dane za jeden rok, a innej – nawet za całe dziesięciolecie milenijne. KRPLOM liczy ponad pół miliona niesymetrycznych plików http: pojedynczy plik może zawierać cały numer gazety, jak również pojedynczy artykuł prasowy lub jedynie spis treści numeru.

Prasę lokalną KRPLOM-u stanowią dzienniki, tygodniki regionalne (tzw. *областные*, np. *Звезда Алтая, Карельская Губерния*), miejskie (np. *Архангельск, Каргополье*) oraz centralne – stołeczne (Sankt-Petersburg, Moskwa: *Аргументы и Факты, Петербургский Час Пик, Российская газета*). Gazety te oświetlają wydarzenia poszczególnych

³ Aktualnie do Rosji dołączone zostały Republika Krym oraz miasto Sewastopol o znaczeniu regionalnym, przy czym państwa, organizacje, m.in. ONZ nie uznają przyłączenia tych terenów do zintegrowanych terytoriów Federacji Rosyjskiej. To samo dotyczy Donieckiej i Ługańskiej Republik.

⁴ Obecnie FR składa się z 89 RFS: 48 regionów, 24 republik, 9 krajów, 3 miast o znaczeniu federacyjnym, 4 okręgów autonomicznych oraz 1 obwodu autonomicznego. Przynależność niektórych RFS do FR jest sporna: uznawane są one jedynie na terenie FR. Jest to związane z polityką państwa, np. militarną agresją Rosji w Ukrainie, wojną informacyjno-propagandową na Donbasie i Ługańszczyźnie, aneksją Krymu. Będziemy sugerować podziałem administracyjnym z czasów ekstrakcji AS dla niniejszego repozytorium do roku 2014.

regionów, miast, wsi, tj. informacje o statusie czysto lokalnym o tematyce społeczno-politycznej (np. polityka, kultura, sport, edukacja, religia, ekologia, wiadomości kryminalne i in.).

Dla potrzeb korpusowych na rok 2006 odszukałem 2302 (*sic!*) periodyki centralne oraz regionalne, mające wersje zarówno drukowane, jak internetowe. Z internetu udało się pobrać jedynie 280 gazet mających odpowiedniki online, czytelną strukturę html stron, tj. „opłacalne” ze względu na wykonalność obróbki czysto technicznej (np. zmiana kodowania standardu czcionki, możliwość wykonania kopii etc.). Parametrem zasadniczym był aktywny serwer określonej gazety w trakcie gromadzenia danych (rok 2006, początek roku 2007). Przed rokiem 2006 uznawano 86 subiektów FR. Wspomniane 280 gazet pochodzi z 71 subiektów – na materiale tekstowym tych gazet utworzono LNS, tj. wstępną listę nowych słów. W niniejszym Repozytorium agnomy słownikowe pochodzą z 235 periodyków, które z kolei pochodzą z 68 subiektów FR. Takie zredukowanie liczby regionów Rosji ma związek z realnym wynikiem ekscerpcji oraz modyfikacją LNS.

Twardy dysk, na którym został ulokowany KRPLM, liczy ponad 650.000 plików tekstowych. Są to zarówno pojedyncze artykuły gazet, jak i całe strony (w terminologii rosyjskiej *газетная полоса*), zawierające od jednego artykułu do zawartości całego numeru (por. tabelę 2).

Przy doborze gazet do KRPLM-u największą uwagę skupiono na prasie lokalnej. Prasa ta koncentruje się na wydarzeniach „bliższych” mieszkańcom poszczególnego regionu, tj. na informacjach czysto lokalnych o tematyce społeczno-politycznej (np. kultura, sport, edukacja, religia, ekologia, listy do redakcji, wiadomości kryminalne i in.). Szczególną uwagę skierowano na dzienniki, publikujące artykuły w stylu drobnych wiadomości⁵ prasowych. W tabeli 2 przedstawiamy reprezentatywny spis periodyków wykorzystanych w niniejszym repozytorium. Spis ten – jawnie i w całości tu podany – uznaję za jeden z najistotniejszych walorów niniejszego opracowania⁶. Ograniczenia korpusowe sprowadzają się do danych liczbowych: z 280 gazet

⁵ Za najlepsze źródło drobnych wiadomości prasowych uznawane (np. Kurcz et al. 1990) są właśnie dzienniki.

⁶ W praktyce badań lingwistycznych nader często podaje się ogólne, by nie powiedzieć: ogólnikowe informacje, całkowicie nieweryfikowalne, dotyczące jedynie wyboru tytułów prasowych poddawanych wrywkowej ekscerpcji.

KRPLOM-u w Repozytorium wykorzystano 235 periodyków z 93 miast mieszczących się w 68 RFS – por. tabelę 2. W tabeli umieszczono następujące dane:

Kolumna 1. Jeden rodzaj danych: lista alfabetyczna RFS.

Kolumna 2. Trzy rodzaje danych:

- nazwa RFS w 2007 r.;
- ogólna liczba wykorzystanych czasopism w IAS;
- ogólna liczba wykorzystanych AS na liście LNS.

Kolumna 3. Jeden rodzaj danych: lista RFS w 2022 r. ze wskazaniem zmian w liczbie i nazwach RFS.

Kolumna 4. Jeden rodzaj danych: nazwa stolicy RFS w 2022 r.

Kolumna 5. Dwa rodzaje danych:

- miasto RFS (w 2007 r.);
- liczba gazet.

Kolumna 6. 2 rodzaje danych:

- nazwa czasopisma;
- liczba AS na liście LNS dla dowolnego czasopisma.

Tabela 2. Dane lingwochronologiczne w RAS: RFS i AS

1	2	3	4	5	6
	Nazwa RFS w 2007 r.; całkowita liczba wykorzystanych czasopism w IAS; ogólna liczba wykorzystanych AS	Lista RFS w 2022 r. ze wskazaniem zmian w numerach i nazwach RFS	Nazwa stolicy RFS w 2022 r.	Miasto RFS (na rok 2007); liczba czasopism	Nazwa periodyku; liczba fragmentów AS w IAS dla każdego periodyku
1.	Алтайский край; 6; 473	Алтайский край	Барнаул	Барнаул; 6	Алтай Daily Review; 131 Ваше дело; 30 Комсомольская правда – Алтай; 21 Молодёжь Алтай; 20 Монитор+ (czasopismo); 240 Свободный курс; 31
2.	Амурская область; 3; 161	Амурская область	Благовещенск	Благовещенск; 2	Амурская правда; 145 Комсомольская правда – Приамурье; 15 ТВС-анонс; 1

1	2	3	4	5	6
3.	Архангельская область; 12; 140	Архангельская область	Архангельск	Архангельск; 6	Архангельск; 11 Архангельская лесная газета; 4 Ведомости Поморья; 1 Волна; 15 Дачная; 2 Двинская правда; 5
				Шенкурск; 1	Важский край; 6
				Каргополь; 1	Каргополье; 3
				Северодвинск; 4	Завалинка; 3 Корабельная сторона; 58 Северный рабочий; 27 Травинка; 5
4.	Астраханская область; 3; 227	Астраханская область	Астрахань	Астрахань; 3	Комсомолец Каспия; 2 Волга; 67 Астраханские ведомости; 158
5.	Белгородская область; 3; 37	Белгородская область	Белгород	Белгород; 2	Житьё-Бытьё; 15 Миссионерское обозрение (czasopismo); 1
				Старый Оскол; 1	Старый Оскол; 21

1	2	3	4	5	6
6.	Брянская область; 2; 27	Брянская область	Брянск	Брянск; 2	Аргументы и Факты – Брянск; 5 Брянское время; 22
7.	Владимирская область; 1; 2	Владимирская область	Владимир	Владимир; 1	Владимирские ведомости; 2
8.	Волгоградская область; 3; 47	Волгоградская область	Волгоград	Волгоград; 3	Волга и Дон; 1 Волгоградская правда; 34 Комсомольская правда – Волгоград; 12
9.	Вологодская область; 4; 126	Вологодская область	Вологда	Вологда; 4	Вологодская неделя; 1 Комсомольская правда – Вологда; 15 Красный Север; 3 Премьер; 107
10.	Воронежская область; 2; 17	Воронежская область	Воронеж	Воронеж; 2	Берег; 14 Подъем (szasorismo); 3
11.	Еврейская автономная область; 0; 0	Еврейская автономная область	Биробиджан	N/A	N/A
12.	Ивановская область; 2; 120	Ивановская область	Иваново	Иваново; 2	ЧастниК°; 119 Хронометр – Иваново; 1

1	2	3	4	5	6
13.	Иркутская область; 1; 36	Иркутская область	Иркутск	Иркутск; 2	Аргументы и факты – Восточная Сибирь; 36
14.	Кабардино- Балкарская Республика; 1; 48	Кабардино- Балкарская Республика	Нальчик	Нальчик; 1	Газета Юга; 48
15.	Карачаево- Черкесская Республика; 0; 0	Карачаево- Черкесская Республика	Черкесск	N/A	N/A
16.	Калининградская область; 2; 83	Калининградская область	Калининград	Калининград; 2	Калининградская правада; 82 Янтарный Караван; 1
17.	Калужская область; 1; 54	Калужская область	Калуга	Калуга; 1	Весть; 54
18.	Камчатская область (до гоку 2006); 1; 63	Камчатский край	Петропавловск- Камчатский	Петропавловск- Камчатский; 1	Камчатское время; 63
19.	Кемеровская область (до 2019); 2; 4	Кемеровская область – Кузбасс	Кемерово	Кемерово; 1 Новокузнецк; 1	Томь; 1 Кузнецкий тракт; 3
20.	Кировская область; 3; 134	Кировская область	Киров	Киров; 3	Комсомольское племья; 1 Вятский край; 128 Вятский наблюдатель; 5

1	2	3	4	5	6
21.	Костромская область; 0; 0	Костромская область	Кострома	N/A	N/A
22.	Краснодарский край; 1; 1	Краснодарский край	Краснодар	Сочи; 1	Курортный проспект; 1
23.	Красноярский край; 1; 1	Красноярский край	Красноярск	Норильск	Заполярная правда; 1
24.	Курганская область; 0; 0	Курганская область	Курган	N/A	N/A
25.	Курская область; 0; 0	Курская область	Курск	N/A	N/A
26.	Ленинградская область; 0; 0	Ленинградская область	Санкт-Петербург, Гатчина	N/A	N/A
27.	Липецкая область; 0; 0	Липецкая область	Липецк	N/A	N/A
28.	Магаданская область; 1; 5	Магаданская область	Магадан	Магадан; 1	Вечерний Магадан; 5

1	2	3	4	5	6
29.	Москва; 28; 2371	Москва	Москва	Москва; 28	<p>Авторевю; 1 Вестник МСПС (czasopismo); 1 Гудок; 33 Завтра; 1373 Молоко (czasopismo); 11 Московская перспектива; 94 Комсомольская правда – Москва; 1 Интерфакс Время; 16 Московские новости; 560 Наш современник (czasopismo); 1 Независимая газета; 4 Новая газета; 2 Новые Известия; 3 Огонёк (czasopismo); 1 Подвиг (czasopismo); 2 Подъём (czasopismo); 28 Полдень (czasopismo); 1 Россия; 31</p>

1	2	3	4	5	6
					<p>Российская Бизнес-газета; 1 Российская газета; 3 Русская Газета; 1 Русская жизнь (czasopismo); 3 Русский курьер; 63 Сети/Network world (czasopismo); 1 Слово (czasopismo); 1 Совершенно секретно; 102 Современник; 1 Хронос (zasób produkcyjny); 29 Элитный персонал; 1</p>

1	2	3	4	5	6
30.	Московская область; 7; 100	Московская область	Москва, Красногорск	Ногинск; 2 Щёлково; 1 Дмитров; 1 Сергиев Посад; 1 Серпухов; 2	Богородские вести; 11 Богородский край; 1 Время; 14 Дмитровский Вестник; 19 Новое Зеркало – Сергиев Посад; 1 Серпуховские вести; 43 Серпуховские новости; 11
31.	Мурманская область; 27; 26	Мурманская область	Мурманск	Мурманск; 1	Мурманский Вестник; 27
32.	Ненецкий автономный округ; 1; 2	Ненецкий автономный округ	Нарьян-Мар	Нарьян-Мар; 1	Известия НАО; 2
33.	Нижегородская область; 3; 155	Нижегородская область	Нижний Новгород	Нижний Новгород; 2 Саров; 1	Курс Н; 52 Нижегородские новости; 1 Городской Курьер; 102

1	2	3	4	5	6
34.	Новгородская область; 3; 5	Новгородская область	Великий Новгород	Великий Новгород; 3	Волхов; 2 Новгородские ведомости; 1 Торговая сторона; 2
35.	Новосибирская область; 5; 622	Новосибирская область	Новосибирск	Новосибирск; 5	Честное слово; 36 Сибирские огни (szasoripistó); 36 Континент Сибирь; 547 Вечерний Новосибирск; 2 Комсомольская правда – Новосибирск; 1
36.	Омская область; 1; 12	Омская область	Омск	Омск; 1	Коммерческие вести; 12
37.	Оренбургская область; 3; 143	Оренбургская область	Оренбург	Оренбург; 2	Яикъ; 1 Вечерний Оренбург; 4
38.	Орловская область; 2; 71	Орловская область	Орёл	Орск; 1 Орёл; 2	Орская Хроника; 138 Просторы России; 69 Орловский Меридиан; 2

1	2	3	4	5	6
39.	Пензенская область; 3; 157	Пензенская область	Пенза	Пенза; 3	Аргументы и факты – Пенза; 30 Молодой ленинец; 115 Московский комсомолец – Пенза; 12
40.	Пермская область (Пермский край – od roku 2005); 10; 1078	Пермский край	Пермь	Пермь; 10	Enter-Профи (czasopismo); 18 Звезда; Звезда-online ; 274 Капитал-Weekly; 65 Комсомольская правда – Пермь; 14 Местное время – Пермь; 44 Пермь; 44 Новости ПГУ; 3 Новый компаньон; 366 Пермские новости; 263 Пермский университет; 27 Хронометр – Пермь; 4
41.	Приморский край; 1; 74	Приморский край	Владивосток	Владивосток; 1	Арсеньевские вести; 74

1	2	3	4	5	6
42.	Псковская область; 2; 229	Псковская область	Псков	Псков; 2	Псковская губерния; 219 Псковская Правда; 10
43.	Республика Адыгея; 0; 0	Республика Адыгея	Майкоп	N/A	N/A
44.	Республика Алтай; 2; 26	Республика Алтай	Горно-Алтайск	Горно-Алтайск	Постскриптум; 24 Правда; 2
45.	Республика Башкортостан; 6; 152	Республика Башкортостан	Уфа	Уфа; 6	Бельские просторы (czasopismo); 106 Уфимский меридиан; 1 Уфимские новости; 1 Йэшлек; Молодость ; 36 Версия Башкортостана; 4 Московский комсомолец – Уфа; 4
46.	Республика Бурятия; 0; 0	Республика Бурятия	Улан-Удэ	N/A	N/A
47.	Республика Дагестан; 1; 3	Республика Дагестан	Махачкала	Махачкала; 1	Новое дело; 3
47.	Республика Ингушетия; 1; 52	Республика Ингушетия	Магас	Назрань	Сердало; 52

1	2	3	4	5	6
49.	Республика Калмыкия; 0; 0	Республика Калмыкия	Элиста	N/A	N/A
50.	Республика Карелия; 6; 76	Республика Карелия	Петрозаводск	Петрозаводск; 6	Площадь Свободы; 1 Петрозаводск; 4 ТВР-Панорама; 3 Моя газета; 4 Лицей; 9 Карельская Губерния; 55
51,	Республика Коми; 5; 301	Республика Коми	Сыктывкар	Сыктывкар; 4	Молодежь Севера; 141 Зырянская жизнь; 66 Комсомольская правда – Сыктывкар; 39 Verbum; 7 НЭП+С; 48
52.	Республика Марий Эл; 2; 67	Республика Марий Эл	Йошкар-Ола	Ухта; 1 Йошкар-Ола; 2	Марийская правда; 1 Московский комсомолец – Марий Эл; 66
53.	Республика Мордовия; 2; 30	Республика Мордовия	Саранск	Саранск; 1 Ардатов; 1	Вечерний Саранск; 7 Маяк; 23

1	2	3	4	5	6
54.	Республика Саха (Якутия); 2; 56	Республика Саха (Якутия)	Якутск	Якутск; 2	Якутия; 13 Илкэн; 34
55.	Республика Северная Осетия – Алания; 1; 2	Республика Северная Осетия – Алания	Владикавказ	Нерюнгри; 1 Владикавказ; 1	Индустрия Севера; 9 Дарьал (szasorismo); 2
56.	Республика Татарстан; 3; 3	Республика Татарстан	Казань	Казань; 2	Вечерняя Казань; 1 Республика Татарстан; 1
57.	Республика Тыва; 0; 0	Республика Тыва	Кызыл	Набережные Челны; 1 N/A	Челны ЛТД; 1 N/A
58.	Республика Хакасия; 0; 0	Республика Хакасия	Абакан	N/A	N/A
59.	Ростовская область; 5; 208	Ростовская область	Ростов-на-Дону	Ростов-на-Дону; 4	Приазовский край/Стадион; 1 Комсомольская правда – Ростов-на-Дону; 51 Город N; 142 Газета Дона; 1
				Волгодонск; 1	Вечерний Волгодонск; 13

1	2	3	4	5	6
60.	Рязанская область; 1; 7	Рязанская область	Рязань	Рязань; 1	Рязанские ведомости; 7
61.	Самарская область; 7; 185	Самарская область	Самара	Самара; 4	Самарская газета; 1 Ваш выбор; 13 Благовест; 49 Аргументы и факты – Самара; 48
				Сызрань; 1	Гильдия; 2
				Тольятти; 2	Тольяттинское обозрение; 59 Площадь Свободы; 13
62.	Санкт-Петербург; 3; 55	Санкт-Петербург	Санкт- Петербург	Санкт- Петербург; 3	Петербургский Час- Пик; 10 Октябрьская магистраль; 44 Невское время; 1
63.	Саратовская область; 5; 270	Саратовская область	Саратов	Саратов; 5	Саратовский Арбат; 2 Саратовские вести; 93 Неделя области; 79 Богородские вести; 1 Богатеи; 95
64.	Сахалинская область; 2; 3	Сахалинская область	Южно- Сахалинск	Южно- Сахалинск	Наши острова; 1 Сахалинское время; 2

1	2	3	4	5	6
65.	Свердловская область; 6; 143	Свердловская область	Екатеринбург	Екатеринбург; 6	Деловой квартал; 64 Деловой круг; 29 Депутатский круг; 5 Известия Уральского государственного университета (szasorismo); 1 Пенсионер; 17 Тайна жизни; 26
66.	Смоленская область; 2; 7	Смоленская область	Смоленск	Смоленск; 2	Рабочий путь; 6 Никольское кольцо – Смоленск; 1
67.	Ставропольский край; 0; 0	Ставропольский край	Ставрополь	N/A	N/A
68.	Тамбовская область; 3; 3	Тамбовская область	Тамбов	Тамбов; 2	Тамбовский курьер; 1 Тамбовский меридиан; 1
69.	Тверская область; 2; 60	Тверская область	Тверь	Мичуринск; 1 Тверь; 2	Мичуринская правда; 1 Тверские ведомости; 5 Вече Твери; 55
70.	Томская область; 0; 0	Томская область	Томск	Томск	Красное знамя; 15

1	2	3	4	5	6
71.	Тульская область; 2; 44	Тульская область	Тула	Тула; 2	Молодой коммунар; 7 Комсомольская правда – Тула; 37
72.	Тюменская область; 2; 58	Тюменская область	Тюмень	Тюмень; 2	Тюменские известия; 34 Тюменские ведомости; 24
73.	Удмуртская Республика; 0; 0	Удмуртская Республика	Ижевск	N/A	N/A
74.	Ульяновская область; 6; 30	Ульяновская область	Ульяновск	Ульяновск; 6	Деловое обозрение; 9 Аргументы и факты – Ульяновск; 3 Ульяновск сегодня; 1 Ульяновский меридиан; 3 Аргументы и факты – Ульяновск; 3 Вестник – Ульяновск; 13
75.	Хабаровский край; 1; 3	Хабаровский край	Хабаровск	Хабаровск; 1	Амурский Меридиан; 3

1	76.	<p>2 Ханты-Мансийский автономный округ; Ханты-Мансийский автономный округ – Югра (od roku 2003); 9; 786</p>	<p>3 Ханты-Мансийский автономный округ – Югра</p>	<p>4 Ханты-Мансийск</p>	<p>5 Сургут; 6</p>	<p>6 Вестник – Сургут; 299 Вестник; приложение Синица – Сургут; 1 Контакт; 19 Московский комсомолец – Югра; 65 Сибирский газетчик; 9 Сургутская трибуна; 297</p>
				Когалым; 1		Когалымские вести; 3
				Нижневартовск; 1		Местное время; 92
				Урай; 1		Знамя; 1

1	2	3	4	5	6
77.	Челябинская область; 11; 297	Челябинская область	Челябинск	Челябинск; 9	Южноуральская панорама; 45 Челябинский рабочий; 35 Лидер; 1 Златоустовский рабочий; 13 Деловой Урал; 18 Городской Дилижанс; 37 Вечерний Челябинск; 20 Бизнес-класс; 3 Аргументы и Факты – Челябинск; 5
				Миасс; 1	Миасский рабочий; 118
				Магнитогорск; 1	Магнитогорский рабочий; 2
78.	Чеченская Республика (od 2003); 0; 0	Чеченская Республика	Грозный	N/A	N/A

1	2	3	4	5	6
79.	Читинская область; 3	Забайкальский край	Чита	Чита; 3	Экстра реклама; 49 Забайкальский рабочий; 27 Забайкальская новь; 8
80.	Чувашская республика – Чувашия; 1; 13	Чувашская Республика — Чувашия	Чебоксары	Чебоксары; 1	Советская Чувашия; 13
81.	Ямало-Ненецкий автономный округ; 1; 100	Ямало-Ненецкий автономный округ	Салехард	Ноябрьск; 1	Слово нефтяника; 100
82.	Ярославская область; 4; 37	Ярославская область	Ярославль	Ярославль; 3	Екклесиаст; 19 Золотое кольцо; 4 Юность; 10
83.	N/A; 0; 0	Севастополь	Севастополь	Рыбинск; 1	Рыбинская среда; 4
84.	Чукотский автономный округ	Чукотский Автономный округ	Анадырь	N/A	N/A
85.	N/A; 0; 0	Республика Крым	Симферополь	N/A	N/A
86.	Таймырский (Долгано-Ненецкий) автономный округ (до 2005); 0; 0	N/A (obecnie jest to: Таймырский Долгано-Ненецкий район Красноярского края)	Таймыр	N/A	N/A

1	2	3	4	5	6
87.	Агинский Бурятский автономный округ (до 2004); 0; 0	N/A (obecnie jest to Агинский Бурятский округ в составе Забайкальского края)	Агинское	N/A	N/A
88.	Коми-Пермяцкий автономный округ (до гоку 2004); 0; 0	N/A (obecnie jest to: Коми-Пермяцкий округ Пермского края).	Кудымкар	N/A	N/A
89.	Эвенкийский автономный округ (до 2007); 0; 0	N/A (obecnie jest to: Эвенкийский район Красноярского края).	Тура	N/A	N/A
90.	Усть-Ордынский Бурятский автономный округ (до 2006); 0; 0	N/A (obecnie jest to: Усть-Ордынский Бурятский округ Иркутской области)	Усть-Ордынский	N/A	N/A
91.	Корякский автономный округ (до 2006); 0; 0	N/A (obecnie jest to: Корякский округ Камчатского края).	Палана	N/A	N/A

1	2	3	4	5	6
92.	N/A; 0; 0	<p>* Республика Крым (od roku 2014)</p> <p>* jednostka administracyjna nie uznawana przez większość społeczności międzynarodowej ONZ</p>	Симферополь	N/A	N/A
93.	N/A; 0; 0	<p>** Донецкая Народная Республика (od 2022)</p> <p>** jednostka administracyjna nie uznawana przez większość społeczności międzynarodowej ONZ</p>	Донецк	N/A	N/A

1	2	3	4	5	6
94.	N/A; 0; 0	<p>*** Луганская Народная Республика (od 2022)</p> <p>*** jednostka administracyjna nie uznawana przez większość społeczności międzynarodowej ONZ</p>	Луганск	N/A	N/A

Układ danych w tabeli umożliwia niezależne obliczenia statystyczne. Zasięg geograficzny słów IAS to praktycznie cała mapa Federacji Rosyjskiej:



Rysunek 2. Lokalizacje wystąpień słów IAS na mapie regionów [oprac. własne]

Nie podaję adresów internetowych ze względu na ogromną serię modyfikacji o charakterze czysto historycznym: od 2007 roku wiele gazet po prostu przestało istnieć, periodyki pozmieniali swoje adresy bądź nawet nazwy. Szczegółowe informacje dotyczące danych adresowych oraz ilości zasobów znaleźć można we dysertacji Fedorushkov 2008 (rozdział pt. *Prasa regionalna*, 80–95).

IV. Metody badawcze

Choć celu niniejszego Wstępu do repozytorium nie stanowi opis metod badawczych i – ogólnie – nie jest to monografia naukowa poświęcona metodom i technikom uzyskiwania w sposób zautomatyzowany neologizmów z korpusu tekstów, pragnąłbym naszkicować podstawowe zagadnienia pozwalające wyeliminować niektóre wątpliwości w związku z podejrzeniem o niereprezentatywność danych językowych.

Najogólniej rzecz ujmując, zakres metod i technik w badaniu można podzielić na dwa zasadnicze etapy: zbieranie AS oraz komponowanie RAS. Zarówno pierwszy, jak i drugi etap nierozzerwalnie związane są z równoległym wglądem do danych korpusowych. W związku z tym w sposób szczegółowy nie zostaną omówione następujące zagadnienia:

- kwerendy korpusowe;
- narzędzia wykorzystywane do analizy danych i budowy korpusu oraz ważne szczegóły dotyczące zaangażowanych procesów technicznych;
- oprogramowanie, algorytmy, metody analityczne, które zostały wykorzystane do wyodrębniania, przetwarzania i organizowania AS;
- motywy doboru słów;
- kwerendy leksykograficzne i weryfikacja obecności wyrazów;
- motywy eliminacji słów: czyszczenie danych, normalizacja i obsługa odmian dialektycznych,
- zapewnienie dokładności i użyteczności repozytorium;
- wyzwania i ograniczenia metodyczne związane z gromadzeniem i weryfikacją ogromnej ilości danych z gazet regionalnych;
- reprezentatywność danych oraz przydatność danych RAS dla badań w zakresie tendencji morfologicznych.

Wszystkie te zagadnienia są szczegółowo omówione w mojej dysertacji. Ograniczam się jedynie do najważniejszych punktów metody badawczej. W następnych rozdziałach Wstępu dodatkowo można się zapoznać z kryteriami weryfikacji leksykograficznej oraz motywami doboru AS do RAS.

Interesującą nas metodę badawczą można określić jako algorytmizowany szereg czynności ekscerpcyjnych oraz weryfikacyjnych. Ścieżkę

dotarcia do całej listy AS można przedstawić na przykładzie etapów selekcji ekscerpcyjnych i weryfikacyjnych odwzorowanych w tabeli 3, w której dokonuje się modyfikacji listy słów korpusu (LSK) do indeksu AS zawartego w niniejszym RAS. Dodatkowy komentarz do poszczególnych etapów (uzyskanych list optymalizowanych do ostatecznej listy AS) jest umieszczony pod tabelą 3. Uzyskanie określonych list jest jednoznaczne z zakończeniem poszczególnych najważniejszych etapów.

Tabela 3. Ścieżka dotarcia do listy AS

LISTA SKRÓTÓW	OPIS LISTY
LSK	lista słów korpusu
LSKA	lista słów korpusu z danymi morfoanalizy
LNS(AK)	„lista nowych słów” + kody gramatyczne (gramemowe)
LNS	„lista nowych słów”
LNS(K)	kontekstowa lista słów
LNS(L)	lista słów poddanych lematyzacji
AS(BVD)	lista AS przed weryfikacją w słownikach rosyjskich
AS(AVD)	lista AS po weryfikacji w słownikach rosyjskich.
AS(BVC)	lista AS przed weryfikacją w kontekście i danych uzupełniających
AS(AVC) = IAS	lista AS po weryfikacji w kontekście oraz danych uzupełniających

LSK. Uzyskanie wszystkich słów korpusu KRPLOM. Narzędzie do uzyskania – to indeksator i segmentator dtSearch Desktop (dtsearch.com) umożliwiający podział sekwencji znaków na tokeny od spacji do spacji.

LSKA. Poddanie słów korpusu analizie morfologicznej oraz uzyskanie wyników takiej analizy z najważniejszymi informacjami: lemma słowa, kod dla klastru kategorii gramatycznych, informacja o obecności

słów w słowniku Zaliznyak 1987. I etap kwerend korpusowych. Użycie technologii AOT dla analizy morfologicznej języka rosyjskiego (aot.ru).

LNS(AK). Uzyskanie surowej listy potencjalnych AS z kodami gramatycznymi, pozwalającymi ekscerpować dokładne słowoformy wszystkich części mowy. Lista zawiera homografy. Motywy doboru przyszłych AS opisany jest w rozdziale VII. *Kryteria ekscerpcyjne* niniejszego Wstępu. Wykorzystanie składni wyrażen regularnych w filtrach ekscerpcyjnych w edytorach tekstowych.

LNS. Lista potencjalnych AS pozbawiona homografów oraz kodów gramatycznych.

LNS(K). Wybór słów, które są poddawane próbnej kontekstualizacji. Sprawdzenie jakości informacyjnej kontekstów. II etap kwerend korpusowych.

LNS(L). Lista potencjalnych AS zlematyzowanych do form podstawowych (np. bezokolicznik dla czasownika, mianownik liczby pojedynczej dla rzeczownika itp.).

AS(BVD). Lista AS przed weryfikacją w zbiorze słów słowników języka rosyjskiego.

AS(AVD). Lista AS uszczuplona o wyrazy wykryte w zbiorze słów słowników języka rosyjskiego. Weryfikacja obecności słów odbywała się drogą zautomatyzowaną, jak i manualną. Kwerendy słownikowe. Opis weryfikacji opisany jest w następnym rozdziale.

AS(BVC). Lista AS przed weryfikacją w użyciu kontekstowym. Weryfikacja poprawności słów oraz pierwsze obserwacje kontekstowe.

AS(AVC). Lista AS po weryfikacji kontekstowej z dołączonymi metadanymi typu tytuł gazety, jej numer oraz datacja. III etap kwerend korpusowych.

Czynności kontekstualizacyjne oraz obróbka materiału, które oddzielają listę AS(BVC) od listy AS(AVC) zajęły ponad dziesięć lat. Uzyskanie listy DA(AVC) zawierającej ponad 10 tysięcy wierszy z kontekstami umożliwiło utworzenie Repozytorium.

V. Weryfikacja agnonimów słownikowych

Jak już wspomniałem, uzyskanie AS oraz opracowanie każdego wyrazu wiązało się z czynnościami zautomatyzowanymi – z użyciem zaawansowanych narzędzi komputerowych, przede wszystkim analizatora morfologicznego oraz oprogramowania służącego do nawigacji w korpusie tekstów elektronicznych. Uzyskiwanie 10 tys. kontekstów oraz metadanych tych kontekstów w korpusie odbywało się systematycznie i manualnie, tj. bez użycia narzędzi pozwalających np. użyć masowej (tj. dla wskazanej liczby ekscerptów), „hurtowej”, chaotycznej ekscerpcji przypadkowych kontekstów. Analizator morfologiczny (technologia AOT) wykorzystany do ekscerpcji zintegrowany jest ze słownikiem *Грамматический словарь русского языка* (Зализняк 1987), a dokładniej – z danymi gramatycznymi podanymi w tym słowniku dla każdego wyrazu. Jest to trzecia⁷ modernizowana wersja licząca 100 tys. słów. Biblioteki analizatora pozwoliły rozpoznać każdą słowofromę tych słów również w korpusie KRPLOM. Należy wskazać w tym miejscu, że dany słownik bazuje na leksyce czterech (sic!) słowników objaśniających: najstarszego, 4-tomowego słownika Uszakowa (lata 1935–1940), 1-tomowego słownika Ożegowa (1 wydanie – to rok 1947), a także dwóch akademickich dzieł: *Słownika języka rosyjskiego*, tzw. MASu, tj. 4-tomowego „Małego słownika akademickiego” (I wydanie – 1961) oraz wydanego w latach 1950–1965 17-tomowego *Słownika współczesnego literackiego języka rosyjskiego*, tzw. BASu, tj. „Wielkiego słownika akademickiego” (por. Bibliografia).

Przeprowadzone zostały konsultacje, tj. weryfikacje obecności użytych słów w dodatkowych źródłach leksykograficznych. LNS słów została skonfrontowana z listą słów ujętych w zespole słowników – por. *Słowniki weryfikacyjne* w Bibliografii.

W ten sposób skonfrontowano LNS zawierającą (potencjalnie) neologizmy z siatką słów już wcześniej zarejestrowanych – przed rokiem 2006. Należy tu stanowczo zaznaczyć, że niniejszy słownik zawiera

⁷ Obecnie mamy do czynienia już z VI wydaniem słownika. Oficjalna wersja online słownika znajduje się na stronie <https://gramdict.ru/> – WEB: 22.12.2023. Słownik liczy około 110 tys. słów, w tym 8 tys. imion własnych.

jedynie te wyrazy, które nie figurowały właśnie w siatkach wyżej wskazanych źródeł. Wyrazy w IAS nie były weryfikowane w słownikach opublikowanych po 2006 roku, oraz danych innych korpusów, które pojawiły się po wspomnianym roku.

W czasie ukontekstowania LNS, a ostatecznie AS(AVC) niejednokrotnie poddana została aktualizacji, modernizacji oraz uszczupleniom, zamianom i uzupełnieniom. Obecna nazwa techniczna tej listy – to Indeks AS w porządku *a fronte*, tj. układzie alfabetycznym haseł od lewej do prawej strony.

VI. Kontekst oraz definicja

Przed rozpoczęciem opisu struktury słownika chciałbym tu umieścić komentarz na temat istoty całej problematyki: dlaczego na tyle ważnym był zabieg ukontekstowania wyrazów LNS uzyskanej dawno temu – bo aż w 2006 roku? Otóż dany zabieg służył przede wszystkim weryfikacji poprawności gramatycznej oraz potwierdzenia racjonalności użycia: przecież to kontekst czyni ze słowa to, co nazywamy leksemem. Sprawą drugorzędną jest jednak to, czy pozostanie on w systemie podobnych mu dyskretnych bytów „od spacji do spacji” czy ulegnie procesowi agnonimizacji, tj. przejściu z peryferii do strefy, której nie dosięga promień pamięci językowej użytkowników. We wstępie do szóstego wydania swojego słownika Anatolij Zaliznyak pisze:

(...) wśród setek nowych słów, które można znaleźć na stronach bieżącej prasy, wciąż bardzo trudno jest odróżnić te, które pozostaną na stałe w języku, od tych, które okażą się zwykłymi jednodniówkami. Ponadto wiele z tych słów nie jest jeszcze stabilnych pod względem składu fonetycznego, akcentu i użycia. Zadaniem pokrycia tej masy niestabilnego nowego materiału leksykalnego powinny zająć się specjalne słowniki nowych słów i znaczeń, ale nie słowniki gramatyczne. (Zaliznyak 2022 z przedmowy do internetowej wersji słownika). [Tłum. z ros.– Y.F.].

Aby uchwycić tę wspomnianą przez Zaliznyaka niestabilizowaną leksykę potrzebny jest właśnie kontekst. Innymi słowy, podanie faktu, że wyraz już „rozpoczął karierę” w ramach leksykalizacji. Ponadto sam fakt odnalezienia się słowa w kontekście już wzbogaca system semantyczny: wyraz wchodzi w gotowe pola semantyczne, aktualizuje takie pola bądź inicjuje pola również nowe, uplastycznia i urozmaica grawitację między słowami, wzmacnia syntagmatykę, uzupełnia lanki oraz minucje językowe i pojęciowe, wzmacnia możliwości aprioryczne języka, optymalizuje system nawigacji i orientacji użytkownika w czasoprzestrzeni pozajęzykowej – czyli *de facto* to wszystko, co wspomaga i unowocześnia sposób myślenia współczesnego użytkownika języka.

Innym niezaprzeczalnym faktem jest też to, że takie wyrazy – rzecz jasna, nie wszystkie – faktycznie się leksykalizują, osiadają nie tylko na bliższej lub dalszej peryferii uzusu. Część nowych słów ze względu na zapotrzebowanie intelektualne może pozostać jedynie w roli udanego lub nieudanego okazjonalizmu dziennikarskiego lub błyskotliwego autorskiego hapaksu, który jednak nie kontynuuje leksykalizacji⁸, nie służy jako prefabrykat w kolejnych kontekstach. Oprócz racjonalności użycia nowego słowa ważna jest poprawność jego użycia; słowo ma spełniać kryteria odpowiadające ogólnie nomologicznym (tj. prawidłowościowym) zasadom językowym.

RAS nie ma charakteru normatywnego, nie opiera się na kodyfikacji normatywnej, w tym ortograficznej. Zachowana pisownia jest oryginalna. Można go określić jako faktograficzny lub jako idiologiczny, tj. zdarzeniowy, odnoszący się do faktów językowych wspartych informacją czasowo-lokalną. Służy przede wszystkim demonstracji wzrostu aktywności pojawienia się nowych słów w okresie przełomu dwóch tysiącleci na terenie współczesnej Rosji.

Podkreślę, że nie miałem zamiaru obserwować losu poszczególnych słów. W RAS nie są podane informacje, jak długo poszczególne słowa egzystowały, czy się uzualizowały, czy zostały ujęte przez słowniki nowych słów i in. Rzecz oczywista, że tak potężną – opartą na „nielimitowanej” wręcz masie tekstowej – a także leksykograficzną weryfikację warto było przeprowadzić „w pojedynkę”. Ale „w pojedynkę” nie można było zrealizować każdej możliwości badawczej. Niemniej przy następnych edycjach RAS parametry związane z szerszą notacją są przewidziane.

W artykułach hasłowych nie uwzględniono także definicji. Uważam, że definiens warto formułować dla neologizmów uzualizowanych – tj. takich, które osadziły się w setkach, tysiącach lub nawet milionach kontekstów oraz wciąż występują, tj. kontekstualizują się. Weryfikacja obecności w aktualnym uzusie, obserwacja losu poszczególnych haseł, ich semantyzacja oraz wyznaczenie algorytmu, jak i czynności eksplikatycznych mogą być zadaniem w innych badaniach.

⁸ Mówiąc o leksykalizacji, mam na myśli proces permanenty. Zatrzymanie leksykalizacji – to brak pojawiania się słowa w kontekście zbiorowym – tzw. *multikontekście*, tj. tekstach aktualnie generowanych przez poszczególnych użytkowników, w poszczególnych idiolektach – indywidualnych językach użytkowników.

VII. Kryteria ekscerpcyjne

Wyróżniam cztery ważniejsze kryteria dodawania słów, kontekstów oraz metadanych do IAS z późniejszym włączeniem do Repozytorium:

1. Brak obecności AS w zbiorze *Słowników weryfikacyjnych* (por. podrozdział o tej samej nazwie w rozdziale Bibliografia);
2. Dobór kontekstu o wcześniejszej dacie publikacji;
3. Czytelność danych korpusowych;
4. Samowystarczalność kontekstu oraz zawartej w nim informacji;

Zgodnie z kryterium (1) słowo dołączane do IAS nie powinno być obecne we wcześniejszych źródłach leksykograficznych (por. Bibliografia – *Słowniki weryfikacyjne*. Przykładowo udało się wyeliminować następujące wyrazy zarejestrowane w innych słownikach:

- Котелова 1995 (*Словарь новых слов русского языка* (середина 50-х – середина 80-х гг.)): *актировка, виброрейка, детприемник, кинонаблюдение, котлоагрегат, культстан, магазин-клуб, рок-музыкант, рок-сцена, телесмотрение, телеящик, фильмобаза, фотожурналистка*;
- Котелова 1996 (*Словарные материалы... – 1986*): *гипнабельный*;
- Котелова, Денисенко 1996 (*Словарные материалы... – 1987*): *хард-роковый, шоу-группа*;
- Соловьёв 1996 (*Словарные материалы... – 1989*): *киножитие, киножурналистика, киномонстр, общак, пипл, полузапрещенный, снэк-бар, украинофобия, ультрамарафон*;
- Левашов 1996 (*Словарные материалы... – 1998*): *"химифобия", самсебяиздат, телемагнат, ультраранний, шоу-представление, эмбриотоксический*;
- Буцева 2004 (*Словарные материалы... – 1990*): *абсолютка, эротик-шоу*;
- Милековская et al. 1996 (tom III): *сверхдоходы, чоппер*;
- Милековская et al. 1990 (tom II): *фашизоид*;
- Милековская et al. 1988 (tom I): *кок-чай*.

Spośród kontekstów KRPLOM-u zawierających poszukiwane słowo dobierany był kontekst z możliwie najwcześniejszą datą ukazania się w wersji internetowej gazety lokalnej. Dobór ten stanowił kryterium

(2). Spośród kilku potencjalnych ekscerptów wybierałem wcześniejszy, tj. i. oraz iv. – zob. niżej:

- i. Бодро прошел концерт томской группы "Будни лепрозория". Последние пару лет ребята, "засветившиеся" еще в 80-х, играют с **драм-машиной**. В их составе давно уже нет саксофона, впрочем, трубы, гитары и баса вполне достаточно для мощного драйва, инициируемого вокалистом и гитаристом Глебом. <Континент Сибирь №7 (249) 2001>
- ii. После безуспешных попыток найти подходящую ритм-секцию "Шутки" ограничились **драм-машина**. <Псковская губерния №2 (221) 2005>
- iii. Чаще всего там появлялась с наркотиками подручная "**наркобаронессы**" Эльвира, но могла "осчастливить" жаждущего клиента и другая помощница – Елена. <Любимая Газета – Пенза №38 2004>
- iv. Клиентами "**наркобаронессы**" являлись в первую очередь ранее судимые, причем по статье о незаконном обороте наркотиков, жители города Никольска и Никольского района. <Молодой ленинец №20 (6961) 2000>

Kryterium (2) dotyczy lokalizacji czasowej słów IAS i wiąże się z czynnościami doboru kontekstu powstałego w punkcie czasowo najwcześniejszym. Zdaję sobie sprawę, że proponowana w IAS datacja związana jest *de facto* z czynnością pierwszego uzyskania datacji (tj. Datacja 1.0) jedynie w obrębie KRPLOM-u. Uznaję również, że klarowność oraz dokładność danego kryterium mogła zostać „przysłonięta” a) subiektywnym wyborem kontekstu, b) zwyczajnym brakiem innego kontekstu, c) brakiem datacji na stronie www artykułu, który powstał wcześniej. Mimo tych punktów uznaję, że metodycznie przestrzegałem zasady wynikającej z danego kryterium – dobierać kontekst użycia (o ile oczywiście taki wybór miał miejsce) w najwcześniejszym punkcie czasowym.

Dlaczego jednak możliwe oraz wskazane są czynności związane z modyfikacją IAS? Dotyczy to przede wszystkim zmiany datacji, tj. możliwego przesunięcia (jedynie „dolnego” obniżenia) granic Datacji 1.0. Pierwsza pełna redacacja, czyli Datacja 2.0. może być dziełem jedynie kolektywnym oraz kolejnym etapem weryfikacji lingwo-

ochronologicznej agnonimów słownikowych znajdujących się w IAS. Tym samym AS dołączyłyby do wspólnej bazy łączącej poszczególne punkty na hipotetycznej mapie datacji również innych słów. Praktycznie rzecz ujmując, pierwsza przewidywana czynność podczas redatacji to sprawdzenie słowa w innych zasobach tekstowych oraz obserwacja losu słowa – tj. występowania w kontekstach do określonego momentu od zarejestrowanego momentu *ab inculnabūlis*. Niemniej już obecna wersja RAS (z datacją 1.0 przeprowadzonej w przestrzeni tekstów jedynie samego korpusu KRPLOM) może posłużyć do szerokich badań lingwistycznych, w szczególności neograficznych/neologicznych, leksykologicznych, morfologicznych lub lingwochronologicznych.

Jako osobne kryterium (3) wyróżniam czytelność danych korpusowych. Mam tu na myśli poprawne czytelne kodowanie stron protokołu HTTP zawierające artykuły gazet. W okresie powstawania KRPLOMu najbardziej popularnymi kodowaniami były WIN-1251 oraz KOI8-R, które w określonych sytuacjach (np. w przypadku sprzeczności kodowań) uniemożliwiały poprawny odczyt danych. Kodowanie Unicode (UTF-8) nie było wówczas standardową normą kodowania symboli „grażdanki” na stronach www, tak jak ma to miejsce obecnie. Podczas czynności ekscerpcyjnych największy nacisk został położony na przyrodę zarówno gramatyczną wyrazu, jak na jego unikatowe właściwości diakrytyczne. Można zatem stwierdzić, że użyta metoda to metoda ekscerpcji diakrytycznej, uwzględniająca różne rodzaje kodowań. Jeden z problemów technicznych polegał właśnie na ujednoznacznieniu kodowania nawet całych archiwów.

Podstrona internetowa z artykułem gazety miała zawierać datę jej publikacji, przynajmniej rok lub numer. Figurowanie na stronie www numeru gazety (tj. ekwiwalentu drukowanego) umożliwiło rozpoczęcie "śledztwa leksykograficznego" – dotarcia do prawidłowej datacji ekscerptu.

Samowystarczalność kontekstu i jego informacyjność jako kryterium (4) oznacza, że spośród możliwych kontekstów użycia AS dobierany był taki, na podstawie którego można było jedynie intuicyjnie wywnioskować możliwe znaczenie. Znaczenie nowych słów nie do końca jest stabilne i utrwalone. Kontekst składający się z jednego zdania często należało poszerzyć nawet do kilku zdań lub przywołać zdanie z dużo

wcześniej lub później partii tekstu. Ponadto znaczna część RAS – to po prostu wyrazy powstałe na potrzeby chwili. Należy też uwzględnić to, że w przypadku rozpoczęcia prac nad formułowaniem definiensów należałoby zestawiać znaczenia neonomów ze znaczeniami nowymi, które wybrały istniejącą formę graficzną. Innymi słowy IAS nie zawiera neosemantyzmów – tak, jak na przykład *Rosyjsko-polski słownik innowacji leksykalnych* autorstwa Bożeny Hryniewicz-Adamskich. Tytułem przykładu: dla hasła *валюта* Autorka poza znaczeniem podstawowym "1. waluta (...)" podaje również znaczenie semantycznie innowacyjne – "2. *przen.* nazwa koloru nadwozia samochodu (szarozielony)" (Hryniewicz-Adamskich 2005:16). Neologiczne jest więc tu znaczenie 2., które jednak neonimem (przynajmniej w swojej postaci graficznej) nie jest. Zdajemy sobie sprawę, że w ramach innych badań – na przykład analizy obecności wyrazów w aktualnym (w roku 2024) użyciu należałoby wykonać czynności związane z definiowaniem. Wyrazy z IAS pochodzące z dekady 1996–2006, które pozostały w uzusie, nazwalibyśmy wówczas... *neologizmami*. Dlatego że przydzielenie statusu neologizmu wyrazowi następuje, gdy wiadomo, że wyraz niegdyś nowy wciąż jest używany. Tutaj zasadnicze słowo – to "niegdyś", bowiem neologizmy to „słowa, znaczenia słów, idiomy, które istnieją zwykle w pewnym okresie (...)" (Kotelova 1978: 22).

Warto dodać, że przy doborze kontekstów wykorzystana została wyszukiwarka indeksatora korpusu prasy. Indeksator to program, który pozwolił zindeksować, tj. posegregować artykuły gazet w archiwach, oraz same wyrazy zgodnie z wstępnie ustalonym alfabetem (por. niżej). Udało się w ten sposób uzyskać dyskretne byty językowe – słowa jako segmenty znaków alfabetu od spacji do spacji lub znaku interpunkcyjnego. Za znaki alfabetu – poza 66 literami rosyjskimi – przyjęto również cudzośćłów, łącznik oraz 52 litery alfabetu łacińskiego. Wyeksperpowano w ten sposób takie unikatowe dyskretne autonomiczne byty graficzne, jak "sms-ка", "sms", "SMS", SMS, "SMS"ка, "смс"-ка, "SMSка", "смс-ка", "смс", "СМС", SMSка, смс-ка, смс, СМС, "SMS-сообщение", "sms-сообщение", "СМС-сообщение", "смс-сообщение", "SMS"-сообщение, "sms"-сообщение, "СМС"-сообщение, "смс"-сообщение, SMS-сообщение, sms-сообщение, СМС-сообщение, смс-сообщение

– a nawet "эсэмэска" oraz эсэмэска. Niewątpliwie rzuca się w oczy chaotyczność, brak skonwencjonalizowanej sekwencji znaków na oznaczenie wiadomości tekstowej *sms*. Przejawem właśnie nowości wyrazu stanowiło użycie cudzysłowu – czyli oznaki zawahania decyzji dziennikarzy w obrębie użycia bądź nieużycia cudzysłowu. Należy tu wspomnieć, że obecnie przyjęte są dwa obrazy graficzne – *смс* pisane alfabetem rosyjskim oraz *sms*. W swojej książce pt. *Ekscerpcja wyrazów cudzysłowowych z tekstów prasy rosyjskiej* [Fedorushkov 2009] wspominam, że cudzysłów pełni funkcję adaptacyjną podczas utrwalenia wyrazu w uzusie zanim neologizm (a dokładniej: jego brzmienie, obraz akustyczny) przyjmie utwalony obraz graficzny – skonwencjonalizowany standard w grafii i ortografii, tj. osadzi się w normie językowej. Cudzysłów w neologizmie podczas uzualizacji przeważnie znika, a początkowa albo „przejściowa” funkcja cudzysłowu jest adaptacyjna, tj. metametatekstowa. W związku z tym przykłady typu i) oraz ii) były dołączane do IAS:

i) "**авиализинг**"

На это ему посоветовали поработать по вхождению завода в бюджет по строке "**авиализинг**" и за неделю снизить долги по зарплате до двух месяцев. <Неделя области № 43 (108) 2004>

ii) **авиализинг**

Активно сотрудничаем с лизинговыми компаниями России. На сегодняшний день **авиализинг** – спасательный круг отечественного авиапрома. <Капитал-Weekly №14 (513) 2005>

Dlaczego dobrany okres to lata 1996–2006? Dlaczego nie były brane pod uwagę bardziej symetryczne, ścisłe daty, tj. np. *stricte* dziesięciolecie – 1995–2005? Otóż praca nad LNS miała miejsce właśnie w 2006 oraz 2007 roku z zamiarem analizy „ostatnich” 10 lat, a więc 1996–2006/2007. Dlatego dziesięciolecie millenijne nazywane jest dość umownie, choć zasadniczo nawet przekracza okres dekady o rok. Ponadto w pierwszej połowie lat dziewięćdziesiątych wersji internetowych gazet lokalnych, tj. regionalnych było po prostu niewiele. Będzie o tym mowaw kolejnym rozdziale.

VIII. Okres historyczny: dziesięciolecie millenium dla „języka gazet"

Okres, w którym rozwinęła się dyskusja wokół pojęcia języka gazet, to lata 70.–80. XX wieku (por. Novikova 2002: 38). Jaka jest istota i granica tego pojęcia i czy w ogóle istnieje „język gazet"? Opinie na ten temat są niejednoznaczne. Niektórzy autorzy skupiają się na następującym ujęciu danego zjawiska: o języku gazety mówi się w ramach *stylu publicystycznego* (por. Rogova 1975; Panov 1988) lub w ramach tzw. *podstylu publicystycznego* (por. Kozhina 1977: 179). Niektórzy autorzy wypowiadają się o języku gazety jako o kategorii funkcjonalno-stylistycznej języka – czyli w znaczeniu szerszym niż *styl* jako taki (por. Kostomarov 1971: 257, Rozentala 1977: 39). W.A. Aleksejev stawia pod znakiem zapytania istnienie *stylu gazet* i – więcej – gazetowej odmiany stylu publicystyki we współczesnym języku rosyjskim (por. Aleksejev 1973: 9). Autor ten wymienia m.in.: „1. styl gazetowy i dziennikarski; 2. kategoria funkcjonalno-stylistyczna, szersza niż styl w rozumieniu pkt 1; 3. gazetowa wersja stylu dziennikarskiego" (Aleksejev 1973: 9). Niemniej w powyższych i podobnych pracach podkreśla się swoistość języka gazet i konieczność omówienia „języka gazet". Ważna opinia naukowa w tym zakresie należy do badacza W.G. Kostomarowa, który argumentował za istnieniem jedności języka prasy na podstawie zasady polegającej na tym, że systemowość funkcjonowania stylu określa się jednością lingwistycznych i ekstralingwistycznych współczynników oraz wzajemnym stosunkiem jednostek poszczególnych poziomów języka na podstawie ich wspólnego komunikatywnego zadania (por. Kostomarov 1971). Mówiąc o racjonalnym i emocjonalnym oddziaływaniu na czytelnika, Kostomarov bierze za punkt wyjścia zasadę „występowania wymiennego ekspresji i standardu" (ros. *чередования экспрессии и стандарта*) (Kostomarov 1971: 85), która zapewnia „rzetelne przekazywanie czytelnikowi strony merytoryczno-informacyjnej z jej powszedniością, skutecznością, rygiorem, jednoznacznością, intelektualną trafnością oraz strony wpływowo-organizacyjnej z jej ekskluzywnością, emocjonalnością, wręcz sen-

sacyjnością" (Kostomarov 1971: 85) [tłum. – Y.F.]. O.W. Novikowa uważa, że w dużym stopniu ta zasada wyjaśnia się „z historycznego punktu widzenia" (por. Novikova 2002: 38). Tak, na przykład, badaczka N.L. Sinyachkina twierdzi, że to scentralizowany system władzy (w tym – wymóg ścisłości dla komunikatów obracających się w orbicie głównego organu władzy) w latach 70–80 XX wieku okazał wpływ na rotację i wykorzystanie leksyki. Rzeczono zaczynając od lat 20. XX wieku system ten miał wpływ na wykształtowanie się kanonu stylistycznego oraz stylistycznych kliché w języku gazety (por. Sinyachkina 1997: 8).

Przed rozpadem ZSRR, w okresie *pieriestrojki* styl potoczny nie był dopuszczony do języka gazet, w którym dominował styl biurowy. Od roku 1986 rozpoczęła się destandardyzacja leksyki w języku rosyjskim (por. Novikova 2002: 38) i razem z nią rozchwianie norm literackich. Niezbędne w tym miejscu wydaje się wyszczególnienie właśnie historycznych momentów w ewolucyjnych zmianach, które zaszły zarówno w języku narodowym, jak i w języku prasy.

W ostatniej dekadzie informatyczne zasoby internetowe stwarzają silną konkurencję prasie drukowanej w Rosji. Popularność informacyjnych wydawnictw internetowych gwałtownie wzrosła, dlatego obecnie praktycznie każdy regionalny drukowany periodyk w Federacji Rosyjskiej posiada swoją reprezentację internetową. Znaczący wpływ na taki stan rzeczy miała liberalna polityka zapoczątkowana w 1985 roku przez Michaiła Sergejewicza Gorbaczowa (por. Czubiński 1998: 474–493, Vepreva 2005: 11–15), która stworzyła solidne podstawy dla zmiany charakteru przekazu informacji w mediach, zwłaszcza w prasie. Zaawansowanie przeobrażeń sprawiło, że ekipa Gorbaczowa zaczęła zmieniać postawę wobec Zachodu. W rezultacie m.in. w 1987 roku przestano zagłuszać niektóre zachodnie rozgłośnie radiowe nadające programy o Związku Radzieckim, a w miastach zaczęły się pojawiać zachodnie gazety (por. Czubiński 1997: 375). Jednocześnie mediom zalecono porzucenie dotychczasowych form monologowych i wprowadzenie w programach dialogu, a nadto częstsze poruszanie tematów społecznych, ale w sposób wyłączający mentorstwo i pouczanie (por. Gorbaczow 1989: 95–99). Dany stan rzeczy określono mianem *glasnosti*. W rezultacie Związek Radziecki stał się „najbardziej czytającym

krajem świata" (Donaj 2001: 33). „Poluzowana", a następnie ustawowo zniesiona w sierpniu 1990 roku cenzura i wprowadzenie wolności słowa spowodowało, iż w prasie pojawiły się treści niezależne od dotychczasowych zaleceń partyjno-państwowych (por. tamże).

Demokratyzacja w sferze społecznej, politycznej, ekonomicznej sprawiła, że niektóre z gazet będące organami partii lub państwa zostały sprywatyzowane i przeszły na własność redakcji (a w praktyce: przeszły na własność w ręce tzw. kapitału prywatnego), która – poprzez redaktorów naczelnych – sama decydowała o formie i treści artykułów (np. w gazetach „Комсомольская правда", *Известия* itp.); „sugestie" pozacentralne (tj. np. wypływające z nakazów lub zakazów kierowanych w stronę redakcji przez komercyjnego właściciela mediów) pomijam w tym miejscu z powodu braku dokumentacji na ten temat. Z kolei inne periodyki, wydawane w milionowych nakładach, po upadku ZSRR znacząco traciły na liczbie czytelników: m.in. gazeta *Труд*, której nakład w okresie 1986–1991 wynosił 21 mln egzemplarzy jako własność redakcji, osiągnął 1,3 mln nakładu w porównaniu z tygodnikiem moskiewskim „*Аргументы и Факты*" (tzw. *АуФ*). Nakład tej gazety wynosił wówczas 32 mln egzemplarzy sprzedawanych na terenie całego ZSRR. Był to najwyższy na świecie jednorazowy nakład (por. Donaj 2001: 39–40). Po rozpadzie ZSRR tygodnik *АуФ* notował 3 300 tys. egzemplarzy nakładu. Podsumowując, można stwierdzić, że liberalizacja prasy była najwcześniejszą i najszerzą formą przemian w mediach.

Prasa drukowana znalazła się w opozycji do ideologicznie nastawionych mediów telewizyjnych i radiowych (por. Drawicz 1997: 48–62). Jednak przez ostatnie lata popyt na prasę drukowaną znacznie zmalał (jest to raczej tendencja ogólnoświatowa). Rosjanie już nie czytają tak dużo prasy drukowanej jak niegdyś, w czasie choćby przebudowy systemowej na początku lat 90. Jest to związane z tempem życia, z pośpiechem technologicznym oraz wzrastającą liczbą ludzi mających dostęp do internetu, który w wygodny sposób daje możliwość dotarcia do wielu różnorodnych pozycji prasowych, np. do treści nowego numeru oraz informacji archiwalnych. Warto w tym miejscu zaznaczyć, że niektóre gazety udostępniają elektroniczne wersje nowego numeru w całości dopiero po pewnym czasie od momentu pojawienia się w wersji drukowanej.

Dostęp do archiwów, a przede wszystkim archiwów elektronicznych prasy rosyjskiej przed pojawieniem się internetu, był utrudniony. Internet "istnieje" w Rosji od 1990 roku, przy czym zarejestrowana dla Federacji Rosyjskiej domena ru (przykładowo: www.yandex.ru) zaistniała dopiero w 1994 roku. Był to początek globalizacji internetu w Rosji. Z kolei w 1995 roku otwarto internetową stronę www.ug.ru – gazetę *Учительская газета*. Zdarzenie to spowodowało początek rozwoju digitalizacji i cyfryzacji drukowanych periodyków z udostępnieniem danych w internecie. Gazeta ta, być może, stała się przykładem dla innych periodyków rosyjskich. Stworzona została również internetowa gazeta *Газета.py* niemająca odpowiednika drukowanego.

G.N. Trofimova współczesne wydawnictwa sieciowe, a szczególnie środki masowego przekazu (ros. skrót *СМИ*) określa w sposób następujący: электронные, сетевые, онлайнные oraz виртуальные (por. Trofimova 2001b).

W niniejszej pracy uwzględniam jedynie materiał umiejscowiony w sieci (tj. w wersji elektronicznej), a zatem interesuje mnie taki materiał, który – w postaci elektronicznej – ma odpowiednik drukowany. Mając na uwadze przeniesienie centrum ciężkości na wydanie internetowe lub całkowite pozbycie się wydawnictwa drukowanego, muszę zaznaczyć też fakt, że niektóre treści pojawiły się wyłącznie w Internecie (w wersji elektronicznej gazety) już w tamtym okresie. Wydawnictwa periodyków istniejące w postaci „wersja drukowana = wersja online” nie figurowały jako jedyne podczas gromadzenia KRPLOM-u. Zdaję sobie jednak sprawę z możliwych wyjątków z tej sytuacji. Nie uznaję tego za mankament, który wpłynąłby znacząco na ideę ekscerpcji materiału „jedynie drukowanego”: przecież czytelnicy mają prawo wyboru, a treści elektroniczne i drukowane zaczęły funkcjonować w Rosji jako konkurencyjne wobec siebie.

Przed pojawieniem się Internetu w Rosji możliwość zbierania i badania słownictwa prasowego każdego subiekta FR była związana z jednoczesną pracą wielu badaczy. Niespotykana dotychczas możliwość lingwistycznej obserwacji słownictwa języka rosyjskiego, wyekscerpowanego z prasy elektronicznej, została wykorzystana przez różne grupy naukowe, w tym przez kompanię *Integrum*⁹ oraz przez

⁹ Por. strona <http://svc.integrum.ru> – WEB: 12.12.2007).

twórców¹⁰ projektu *Компьютерный корпус текстов русских газет конца XX-ого века*. Twórcy projektu wykorzystali 13 pełnych archiwów począwszy od roku 1994 do 1997 takich gazet, jak *Завтра*, *Правда*, *Литературная газета*, *Новая газета* i in.

W niniejszej pracy piszący te słowa również wykorzystał możliwość lingwistycznej obserwacji słownictwa języka rosyjskiego w udostępnionych zasobach elektronicznych, przy czym dobór wielogatunkowego materiału gazetowego polegał na wyszukaniu i doborze informacji gazetowej w poszczególnych RFS. Ekspansywny wzrost liczby wydawnictw internetowych odzwierciedla liczba tekstów, a zatem popularność socjolektu publicystycznego, więc również słów – w tym neologizmów. Wzrost liczby neologizmów w prasie lokalnej FR unaoczniamy w serii 11 infografik dla 11 lat. Intensyfikacja kolorystyki (od koloru żółtego po fioletowy) w poszczególnym RFS oznacza właśnie ekspansję AS w wymiarze czasowym i geograficznym:

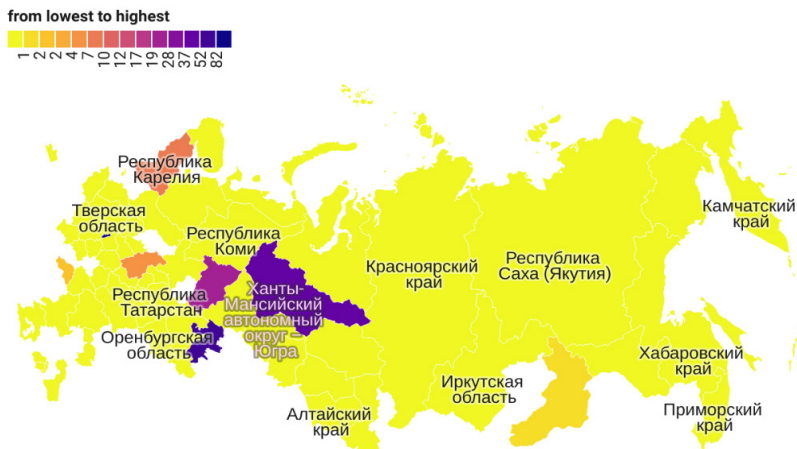


Rysunek 3. Występowanie AS w prasie lokalnej w różnych RFS w 1996 roku [oprac. własne]

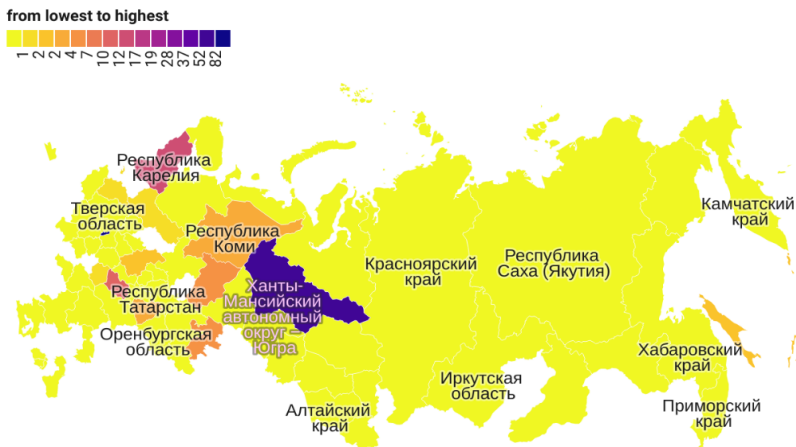
¹⁰ Korpus oraz wersja internetowa korpusu została stworzona w Laboratorium leksykografii ogólnej i komputerowej na Wydziale Filologicznym na Uniwersytecie im. M. Łomonosowa w Moskwie. Autor projektu – A.A. Polikarpov.



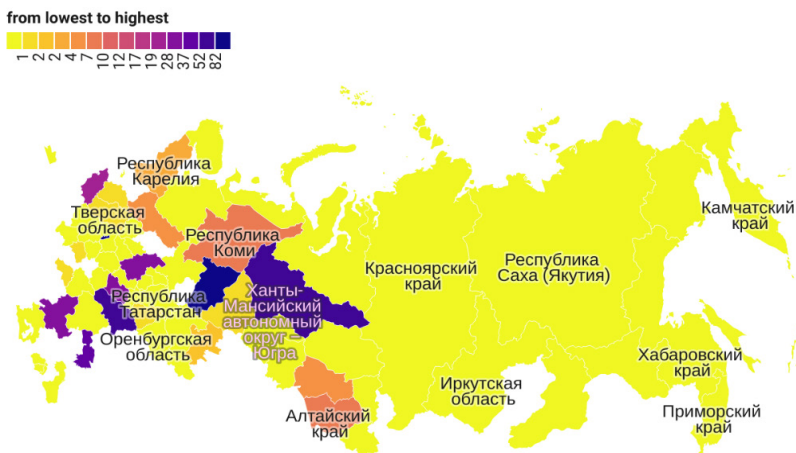
Rysunek 4. Występowanie AS w prasie lokalnej w różnych RFS w 1997 roku [oprac. własne]



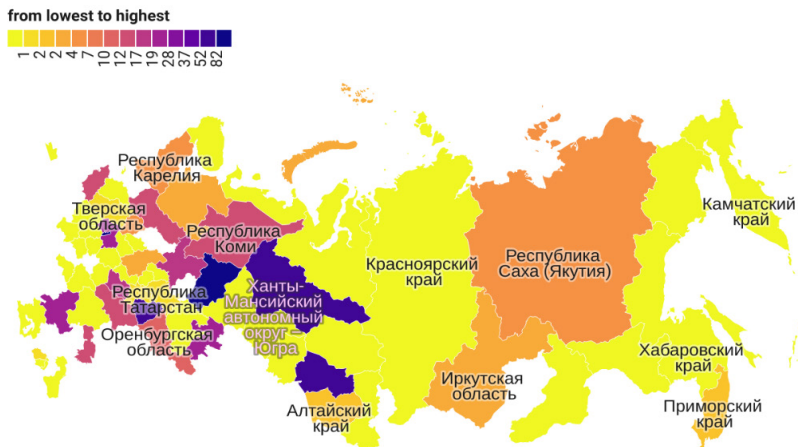
Rysunek 5. Występowanie AS w prasie lokalnej w różnych RFS w 1998 roku [oprac. własne]



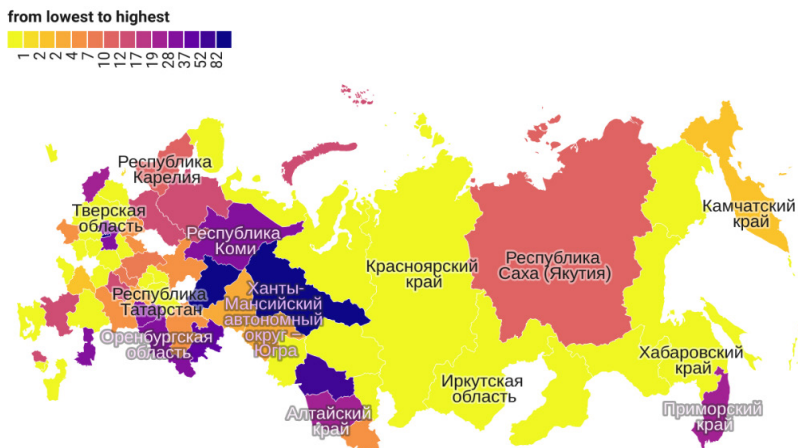
Rysunek 6. Występowanie AS w prasie lokalnej w różnych RFS w 1999 roku [oprac. własne]



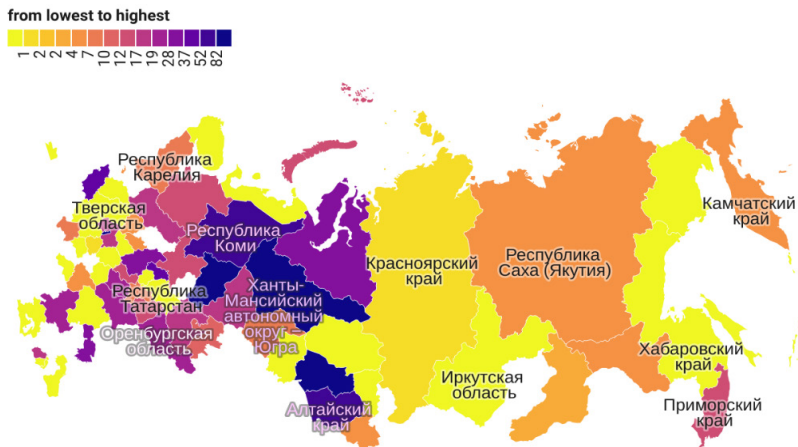
Rysunek 7. Występowanie AS w prasie lokalnej w różnych RFS w 2000 roku [oprac. własne]



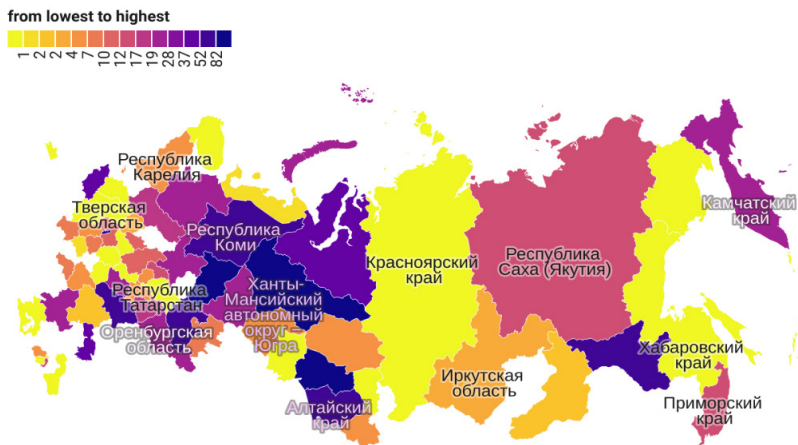
Rysunek 8. Występowanie AS w prasie lokalnej w różnych RFS w 2001 roku [oprac. własne]



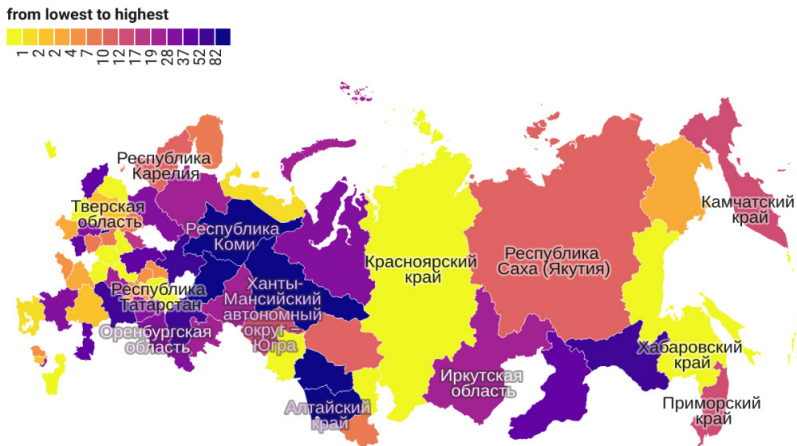
Rysunek 9. Występowanie AS w prasie lokalnej w różnych RFS w 2002 roku [oprac. własne]



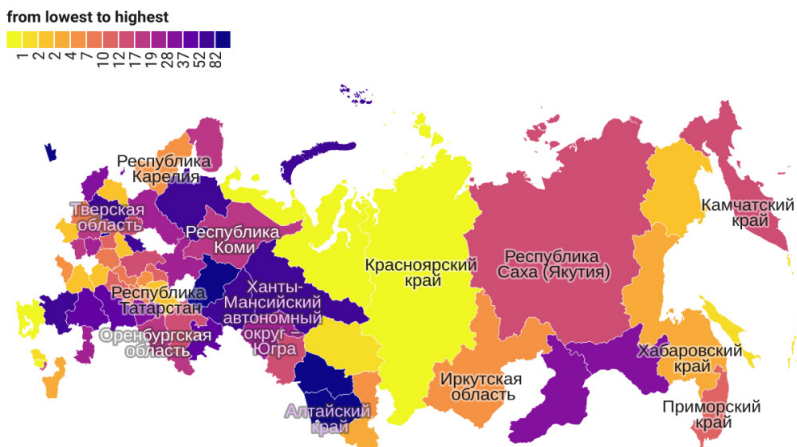
Rysunek 10. Występowanie AS w prasie lokalnej w różnych RFS w 2003 roku [oprac. własne]



Rysunek 11. Występowanie AS w prasie lokalnej w różnych RFS w 1994 roku [oprac. własne]



Rysunek 12. Występowanie AS w prasie lokalnej w różnych RFS w 2005 roku [oprac. własne]



Rysunek 13. Występowanie AS w prasie lokalnej w różnych RFS w 2006 roku [oprac. własne]

Na podstawie powyższych infografik można wnioskować, że wzrost neonimów w ciągu dekady milenijnej skorelowany jest z linią Zachód-Wschód, przy czym szczyt eksplozji (dla większości regionów) to lata 2004–2005. Wydaje się, że podstawowym współczynnikiem wzrostu AS – w pierwszej kolejności neologizmów – jest pojawienie się samych gazet w formacie www, aczkolwiek można także przypuszczać, że wzrost ten związany jest z następującymi czynnikami związanymi z motywacją publiczną utworzenia wersji online:

- optymalizacja gospodarki oraz pozytywne zmiany w zakresie finansowania regionów;
- poszerzenie się działalności Gazpromu i innych korporacji w obszarach trudnodostępnych;
- rozpowszechnienie i ustabilizowanie sieci internetowej;
- swoboda wyrażania prywatnych idei i opinii politycznych;
- dostęp do mediów (prasy, telewizji) pozarosyjskich;
- optymalizacja jakości usług kolejowych i innych systemach transportu, w tym – optymalizacja infrastruktury transportowej, budowa tras federalnych;
- wzrost dochodów, dobrobytu publicznego oraz bogacenie się społeczeństwa z możliwością zakupów publicznych;
- intensywny rozwój technologii komunikacyjnej oraz obliczeniowej;
- pozytywne prospołeczne zmiany polityczne;
- otwarcie na Zachód oraz zachodnie technologie;
- możliwość podróżowania obywateli po świecie;
- „doprowadzenie” internetu do poszczególnych regionów, tj. rozwój infrastruktury ogólnie rozumianych łącz.

Obecna wersja zasobu leksykograficznego RAS umożliwia wykonanie wstępnej prognozy badań: AS na przestrzeni badanego okresu dekadowego mają tendencję do rozwoju w kierunku takich obszarów tematycznych, jak technologie (w tym hi-tech, komputerowe), IT, biznes i marketing, gospodarka, kultura muzyczna. Przypuszczalnie wiąże się to z szeroko rozumianym otwarciem Rosji na Zachód po rozpadzie Związku Radzieckiego. Mediom zalecono porzucenie dotychczasowych form monologicznych i wprowadzenie do dyskursu mowy dialogowej, poruszanie tematów społecznych, ale w sposób eliminujący mentorstwo i pouczanie; liberalizacja prasy była najwcześniejszą i najszerzą

w skali całego państwa formą przemian w mediach. Potwierdza to celna wypowiedź badacza Kostomarowa: „(...) norma literacka i językowa staje się mniej ostateczna i obowiązkowa. Używanie języka rosyjskiego charakteryzuje się obecnie głównym ukierunkowaniem współczesności na wolność, z szalejącą własną wolą, z niedbałością – termin *liberalizacja* jest odpowiedni do określenia tych bardzo szybko rozwijających się procesów” (Kostomarov 1999: 30) [tłum. – Y.F.]. Analiza tematycznego rozwoju słownictwa wymaga jednak osobnych badań.

IX. Możliwości badawcze w zakresie morfologii

W związku z tym Repozytorium może być uznane za źródło reprezentatywne oraz przydatne dla badaczy z zakresu morfologii słowotwórstwa. Mówiąc o tendencjach w rozwoju struktur morfologicznych należy stwierdzić, że słownictwo faktycznie sygnalizuje „wymknięcie się” spod kontroli normy i cenzury w pisowni: *B2B-маркетолог*, *sms-голосование*, *sex-шоп*, *free-чек*), strukturze graficznej i morfologicznej (*слэмить*, *СПИД-"УАЗик"*, *CD/MP3-плеер*), tematycznej i stylistycznej (*спиммер*, *порноюмор*, *поурбаниться*). Szczególną rolę grają tu formanty afiksalne z języków obcych – przede wszystkim z języka angielskiego. I choć aspekt morfologiczny nie znajduje się w zasięgu badawczym niniejszego projektu, w okresie 1999–2003 obserwowalny jest intensywny wzrost użycia zapożyczonych afksoidów oraz radiksoidów o tematyce technologicznej, tj. związane ze wzrostem współczesnych technologii (np. *веб-устройство*, *pdf-файл*, *флэш-карта*, *биоэнерготерапевт*). Z pierwszych obserwacji oraz wstępnych podliczeń wynika, że pod względem budowy morfologicznej neologizmy okresu milenijnego wykazują się wyjątkowo cechami analityzmu, aglutynatywności. Szczególną rolę grają tu formanty inicjalne (prefiksy oraz prefiksoidy) z języków obcych – przede wszystkim angielskiego.

Badaczka Joanna Misturska-Bojanowska w podsumowaniu monografii pt. *Dynamika procesów słowotwórczych w zakresie rzeczowników rosyjskich w drugiej połowie XX wieku* notuje: „Nasza analiza pokazuje, że w sferze produkcji rzeczowników w drugiej połowie XX wieku najbardziej aktywne są następujące procesy: internacjonalizacja, dynamika ilościowego wzrostu leksemów o znaczeniu osobowym, szerokie wykorzystanie nazw własnych jako podstawy bazowe, wzmocnienie kompresji w słowotwórstwie, kolokwializacja i wzrost ekspresywności, wzrost prefiksacji nominalnej, konkurencja i specjalizacja niektórych afiksów (sufiksów i prefiksów), rozprzestrzenianie się synonimii, antonimii i homonimii słowotwórczej, transstopniowe słowotwórstwo, wzrost analityzmu oraz cech aglutynacyjnych” (Misturska-Bojanowska 2013: 137–138). Niektóre z wymienionych procesów można dostrzec „gołym okiem” (tj. bez szczegółowego statystycznego badania) w RAS.

Czeski językoznawca František Trávníček wyróżniał trzy typy hybryd:

- a) wyraz jest utworzony z rodzimej lub przyswojonej podstawy słowotwórczej za pomocą obcego sufiksu,
- b) do rodzimej lub przyswojonej podstawy słowotwórczej dodaje się obcy prefiks,
- c) w obcym wyrazie złożonym jedna część zostaje zastąpiona przez rodzimą" (Trávníček 1948: 30-31 za Mietła 2001: 52).

Nieprzypadkowo cytuję tę osobliwą klasyfikację. Lecz okazuje się, że dana klasyfikacja (opisna przecież dla języka czeskiego, w roku 1948) aktualna jest również dla tendencji rozwoju konstrukcji słowotwórczych rosyjskich wyrazów obserwowanych (wstępnie) w RAS. Innymi słowy, RAS można wykorzystać w badaniach konfrontatywnych, np. transsłowiańskich połączonych z parametrem chronologicznym. Przypuszczam także, że praktycznie każdą klasyfikację dotyczącą tendencji słowotwórczych okresu najnowszego w Rosji (po upadku ZSRR w roku 1991) hipotetycznie można (i warto) potwierdzić, opierając się na zasobie RAS.

Terminy prefiks, prefiksoid (oraz sufiks, sufiksoid) rozszerzam w granicach badania do terminu *formant słowotwórczy*. Przykładowe formanty inicjalne:

sms-, авто-, азото-, айф-, аква-, анти-, аудио-, ауто-, афро-, аэро-, байк-, байто-, баксо-, би-, биг-, бизнес-, билайно-био-, биотехно-, бифи-, блюдо-, бого-, бомбо-, брейк-, бэк-, валюто-, веб-, вебо-, вело-, взаимо-, вибро-, видео-, випо-, власте-, вне-, внутри-, воен-, второ-, вузо-, вшиво-, высоко-, газо-, газонефте-, гейм-, геохимо-, гетеро-, гидро-, гимно-, гип-, гипер-, гипо-, глав-, годо-, гомо-, гос-, гранд-, дам-, дарт-, дач-, де-, дем-, джем-, дип-, диск-, до-, долго-, дон-, дор-, дорого-, дура-, дух-, дымо-, евро-, едрос-, за-, зиц-, идео-, интернет-, инфо-, исламо-, ист-, камне-, кантрик-, квази-, кверху-, кино-, киновидео-, кисло-, книго-, комиксо-, коммуно-, консерво-, контер-, контр-, крано-, кригс-, кубер-, кубино-, кувалдо-, культуру-, курорто-, кучма-, лав-, ланд-, ландро-, ласси-, латино-лево-, лега-, легко-, ледо-, лейко-, лесо-, лже-, ливне-, лизо-, ликеро-, лио-, листо-, лит-, литро-, ложно-, лозо-, лонг-, лонго-, луго-, лыже-, мавро-, майстха-,

макс-, макуло-, мало-, мас-, машино-, мега-, мед-, меди-, медиа-, меж-между-, мезо-, мент-, место-, мета-, метод-, метро-, мех-, мигранто-, микро-, микровидео-, мимо-, мин-, мини-, мино-, мифо-, младо-, много-, моб-, моно-, мор-, москво-, мото-, мотоцикло-, мульт-, мульти-, мульто-, на-, найто-, нано-, нарко-, не-, недо-, нейро-, нео-, нефте-, ново-, ноль-, о-, об-, обл-, обмано-, оген-, одно-, орг-, ортодром-, орф-, осексо-, оф-, пара-, парт-, плей-, пневмо-, поли-, полно-, пост-, про-, проф-, пси-, психо-, психоневро-, радио-, ретро- рок-, рэп-, сверж-, скреб-, смехо-, смс-, софт-, спеле-, стар-, супер-, теле-, телевидео-, телевидеофото-, термо-, тетра-, техно-, фильмо-, фито-, флай-, фото-, фри-, харт-, хэнд-, черно-, чечено-, чисто-, чрез-, шаро-, шело-, шеф-, шино-, шир-, широко-, шифр-, шопо-, шорт-, шот-, шоу-, штраф-, шутц-, щедро-, щито-, эго-, экзо-, эко-, эколесотехно-, эконо-, экс-, экстра-, экстрим-, электро-, эмбо-, эмбрио-, эндо-, эндовидео-, энерго-, энио-, энтеро-, энто-, энхаэл-, эпид-, эро-, эст-, этно-, эхо-, южно-, ява-, языко-, японо- etc.

Tylko wstępna obserwacja IAS wykazuje, że liczba zapożyczonych afiksów i afiksoidów przekracza tysiąc. Przypuszczalnie wiele z nich ma charakter neonimiczny. Niniejszy projekt w jego zrealizowanej formie stwarza perspektywy badawcze w obrębie analizy rozwoju morfologii i morfotaktyki. Problematyka własności morfologicznych – zwłaszcza w obrębie sprecyzowanej klasyfikacji „afiks – afiksoid – półafiks – radiksoid – radiks (rdzeń)” (Minina, Bartkov 1980, Bartkov 2006, Bartkov, Minina 2019) – może stanowić kontynuację badań opisanych w niniejszym projekcie.

W związku z tym należy stwierdzić, że RAS może posłużyć do szerokich badań lingwistycznych, w szczególności leksykologicznych, morfologicznych, neologicznych. Zintegrowane informacje generowane w RAS mogą posłużyć również jako podstawa badań w zakresie historii języka, socjolingwistyki, dziennikarstwa, leksykologii, leksykografii, medioznawstwa. Istotne znaczenie projektu w stosunku do dotychczasowych wyników badań nad unikatowym słownictwem okresu milenijnego będzie takie, że w wyniku badania powstanie największe na świecie repozytorium słów wcześniej nierejestrowanych w języku rosyjskim.

X. Dyskretność słowa oraz niedoskonałości RAS

Neologizm – a jest to centralny obiekt badawczy w niniejszym opisie – reprezentuje nie tylko nowy obraz akustyczny lub graficzny. Jest on przecież powiązany z aktem mentalnym, zostawia po sobie nowy odcisk na mapie obrazów mentalnych skorelowanych z desygnatami – nazwami sytuacji, zjawisk, procesów, przedmiotów, cech, stanów. Zaś wszelka werbalizacja, tj. tworzenie wypowiedzi jest aktem lokucyjnym, który w porównaniu do aktu mentalnego jest precedensem wtórnym. Jest to obraz gotowy, rozpoczynający swoją ewolucję w kontekstach użycia poszczególnych osób, tj. w językach indywidualnych, zwanych też idiolektami. Multilektem – tj. obszarem, w którym kumulują się konteksty użycia – nazywamy wówczas zbiór pojedynczych idiolektów. Możemy tu porównać kumulujące się konteksty użycia tekstowego z kulturą – wspólną identyfikacją, rozpoznawalnością, kodyfikacją i konotacją nowego dyskretnego bytu – neonimu, tj. unikatowej niezarejestrowanej racjonalnie użytej kombinacji liter alfabetu/alfabetów z istniejącym znaczeniem (neonim-„neograf”) lub znaczeniem nowym, niezarejestrowanym (neonim – „neograf” – neosymantyzm, tj. po prostu neologizm). Według antropologa Adama Czyżewskiego, nowy obraz – w tym neologizm – rozumiany jako precedens tworzy przede wszystkim nowo powstałą konwencję, tj. inicjuje kulturę użycia. Generowanie zaś kontekstu użycia – w idiolektach – jest również czymś wtórnym. Czyżewski zaznacza: „Niemożność rozpoznania obrazu wynika (...) nie z braku odpowiedniego kontekstu, a więc społecznie przekazywanych schematów postrzegania, wzorców percepcyjnych, funkcjonujących w umyśle tubylca zaskoczonego zupełnie nową formą przekazu wizualnego, lecz po prostu z braku precedensu” (Czyżewski 2009: 30). Neologizm również reprezentuje sobą punkt odniesienia, który wymaga specyficznego momentu, w którym idiolekty dostosowują się do impulsu kulturowego, do typu lokucji i jednocześnie do konwencji użycia. Taki moment Adam Czyżewski nazywa „krytycznym momentem scalenia idiolektów” dookoła siebie, kiedy przejawia się efekt narodzin wspólnoty „odbioru” i „rozpoznania” (por. tamże) nowego obrazu akustycznego/graficznego – tj. neologizmu. Neologizm

rozumiany jako nowy obraz jest precedensem *per se*, ma status okazjonalny, peryferyjny i agnonimowy. Status ten umożliwia niektórym agnonimom przedostanie się z peryferii do aktywnego i aktualnego centrum uzusu leksykalnego w jego najbardziej żywiołowej i energicznej realizacji – użyciu i dublowaniu się w aktualnie powstających tekstach *hit et nunc*. Neologizm przechodzi do innych idiolektów, tj. figuruje – tak, jak kwant – jednocześnie w kilku lokalizacjach – w kilku punktach multilektu, jakim jest język rozumiany socjolingwistycznie. Należy tu jednak zaznaczyć, powołując się na opinię Czyżewskiego, że pojęcie kultury obejmuje każdy „dowolny, jednostkowy akt mentalny, który w procesie odbioru obrazów wyprzedza moment ujawnienia się presji kontekstu, najogólniej rozumianej wiedzy” (tamże). Narodziny neologizmu w idiolektcie stanowią efekt jednostkowego aktu mentalnego, zaś „zapotrzebowanie w narodzinach” takiego neologizmu jest teoretycznie multilektalne. W określonej mierze narodziny takie są przewidywalne, jeśli taki nowy byt faktycznie wpisuje się we „współczesny” kontekst, ujawniając na zasadzie kontrastu z innymi bytami (słowami) swoją precedensowość. Można też powiedzieć, że neologizm, rozpoczynając dublowanie się w multikonteście, ma w swoim znaczeniu (tj. klastrze semów – semantemie) coś, co wyróżnia go na tle znaczeń innych słów. Można umownie go nazwać suprasemem (tj. wysoce implicytnym „nad”-semem) precedensowości. Hipotetycznie rzecz ujmując, konteksty, które wchłaniają takie znaczenie, z czasem wytrącają taki suprasem, więc neologizm traci najistotniejszą swoją cechę – *novum*. Przewidywalna jest więc też ścieżka rozpowszechnienia, uzualizacji, a więc przyjęcia przez neologizm również określonej struktury (mentalnej → fonetycznej → morfologicznej, semantycznej → walencyjnej) na tle innych (wcześniejszych, tożsamyh czasowo, potencjalnych). Przewidywane powstanie neologizmów może więc mieć walory matematycznie uzasadnionej prognozy. Innymi słowy można przewidzieć, jaką strukturę morfologiczną będzie miał taki nowy obraz akustyczny lub/i graficzny. Jeśli zaś mamy rozwiązanie problemu z określeniem stanu kultury i ogólnie tła mentalnościowego, na którego zapotrzebowanie reaguje też język za pomocą neologizmów, peryfraz etc., to prognoza pojawienia się neologizmu w mowie jest rzeczą teoretycznie możliwą. Dzieje się tak dzięki innemu parametrowi

słowa – dyskretności, czyli podstawowej cesze słowa, która związana jest z autonomią w obecności innych słów, tj. innych – w mniejszym lub większym stopniu – dyskretnych bytów. Ukontekstowanie czyni z wyrazu *leksem*. Następuje faza leksykalizacji, która jest jedną z metamorfoz, napędzająca i modyfikująca nie tylko sam uzus, ale i morfoskładnię, wpływając na potencjał walencyjny poszczególnych bytów dyskretnych, na zwiększenie mas nowych zdań i tekstów, na niezwykle intensywne zmiany w centrum oraz peryferiach uzusu. Ukontekstowanie daje szansę na dalsze życie wyrazu – tj. jego multikontekstową rotację. Słowo jest bytem dyskretnym pod względem morfologicznym i semantycznym, ale prawo ekonomii języka sprawia, że semantycznie słowo nie musi być wcale dyskretne – np. w idomie *czarna owca* nie mamy do czynienia ani z denotatem *owca*, ani z nazwą koloru futra. Dwa słowa tworzą jedno superznaczenie (tzw. znaczenie asumaryczne), które przyjmuje na siebie funkcję dyskretności spośród innych bytów dyskretnych w tekście. Dlatego neologizmy należy obserwować i weryfikować pod względem dyskretności lub jej braku, a także pod względem współwystępowania z sąsiednimi słowami, a dokładniej – ich znaczeniami. Ponadto neologizm, mając nowy obraz graficzno-akustyczny, może nie występować samodzielnie – tylko zawsze z „sąsiadem”. To dlatego RAS ma zasadniczą wadliwość – nieuwzględnienie parametru dyskretności semantycznej: nie obejmuje ani kolokacji, ani idiomów z takim neonimicznym „sąsiadem”. Na przykład kolokacja *лыжероллерный круг* z AS *лыжероллерный* w artykule:

лыжероллерный

Сегодня гордость стадиона – это освещенная лыжная трасса и наполовину построенный **лыжероллерный** круг.
<Архангельск №(47) 2006>

Inne możliwe kolokacje to *попасть в лонглист [конкурса]*, *попасть в шортлист [конкурса]*, *войти в лонг-лист*, *лонг-лист этого года*, *дебютный лонгплей*, *онлайновый музыкальный магазин*, *релиз DVD*, *дебютный лонг-плей*. Wynik: obok neologizmowi może znajdować się „lubiący się z nim” inny wyraz. Taka zbitka wyrazów może reprezentować potencjalnie nowy *frazem*, w którym jeden z komponentów jest neonimem. Kontekstualizacja umożliwia

wyszukiwanie frazemów oraz potencjalnych frazemów (PF), ale także *nowych* frazemów (NF), tj. połączeń wyrazów, które wspólnie wkraczają na ścieżkę frazematyizacji. Jeden pojedynczy neologizm stwarza swój ekosystem przyczyniający się do powstania jeszcze czegoś nowego. Możliwe, że działa tu fenomen koegzystencji (symbiozy) dwóch sieci/dwóch systemów. Porównać to można do grzybów i grzybni (łac. *mycelium*): grzybiarze wiedzą, że jeśli widzi się muchomora, należy szukać w pobliżu innego muchomora albo też... prawdziwka (i odwrotnie). W tekście językowym może również działać zasada koegzystencji dwóch sieci semiotycznych: jeśli znalazłeś neologizm, szukaj w najbliższym otoczeniu kontekstowym innego neologizmu lub coś podobnie nietypowego lub rzadkiego, np. unikatowego pod względem grafii (neonim), nietypowego znaczenia (neosemantyzm, neologizm), konotacji stanowiącej precedens (neologizm), coś nietypowego, wymknięcie się spod konwencji stylu, pisowni, pragmatyki, konotacji etc. Mając to wszystko na uwadze, przypuszczam, że wysoka liczba ukontekstowanych unikatów w określonym zbiorze tekstowym umożliwia prognozę statystyczną liczby innych (albo tych samych) unikatów w innym podobnym zbiorze. Uwzględnienie parametrów kontekstowych możliwe jest przy zastosowaniu sieci neuronowych. Innymi słowy, można „nastróić” sieć na automatyczne wyszukiwanie „nowego/nietypowego” – z pominięciem ekscerpacji manualnej lub automatycznej. Współczynnik kontekstowy współwystępowania neonimów może teoretycznie stać się cechą wektorową sieci neuronowej. Dlatego kontekst występowania uznajemy za najważniejszy parametr dla agnominów słownikowych. W związku z tym RAS określamy jako repozytorium kontekstowe, a nie datacyjne lub geograficzne. Kontekstualizacja wyrazu jest ważna również dlatego, że kontekst w dużej mierze umożliwia zorientowanie się, jaka jest przynależność tematyczna wyrazu, a wręcz jakie znaczenie ma leksem. Pomaga uzmysłwić sposób, w jaki wyraz powstaje dla konkretnego kontekstu, tj. pokonując już pierwszą barierę ku możliwej okazjonalizacji – przejście ze strefy niebytu do strefy egzystencji obok innych jednostek leksykalnych. Dziennikarz, inny twórca wyrazu lub osoba powtarzająca nowo usłyszany wyraz, generując lub regenerując taki obraz akustyczny lub graficzny, nie zniekształca go, przewidując wspólną, powszechnie adaptującą się nazwę dla nieznanego lub słabo

zleksykalizowanego dotąd graficzno-akustycznego referentu potencjalnej nazwy. Dotyczy to np. neologizmu *евросантехника*:

евросантехника

Встретили нас у поезда, привезли в санаторий, водитель проводил к администратору, и за считанные минуты мы были в номере. Импортная мебель, телевизор, холодильник, телефон, **евросантехника**, пластиковые окна, паркетный пол, покрытый коврами, чистота и порядок. <Контакт №11(423) 1999>

Stopień „trafności”, aktualności użycia kontekstowego okazjonalizmu sygnalizuje jednocześnie możliwą prognozę dalszego użycia. Tworzy się wówczas gestalt słowno-pojęciowy, kod, konwencja dla przyszłych użyć w multilektcie, tj. zbiorze tekstów generowanych w zbiorze indywidualnych lektów – idiolektów. W powtórzonym wyrazie ma odbicie swoiste „usolidnienie”, rutynizacja (Bierwiazczonek 2006: 448), jako że używa się prefiksu o konotacji pozytywnej: nazwa produktu, towaru, usługi z prefiksem *euro* – zwłaszcza w marketingu – korelowana jest konotacją „coś solidnego, wzorcowego, szlachetnego, wartego naśladowania”. Powielany jest w ten sposób model stereotypu (Bartmiński 1980: 13–14) w obrębie mody, nowoczesności, zaś produkt, towar lub usługa są solidne oraz dobrej jakości: np. *евровложение, евроремонт, евроотделка*. Choć obok konotacji pozytywnych (Kobozeva 2000) można wyróżnić konotacje nacechowane: *еврозависимый, евровзятка, евробюрократия*. Ciekawostką jest to, że ciąg *евросантехник* w omawianym okresie faktycznie zaczął aktywnie funkcjonować; pojawiła się również nazwa własna, wykorzystywana przez firmy specjalistyczne. Akustyczny obraz nazwy denotatu szybciej się rutynizuje w mowie w porównaniu z wahaniami zapisu graficznego (por. powyższe przykłady z wyrazem *sms*). Przypuszczam zatem, że rzadziej może mieć miejsce labilność fonologiczna takiego obrazu, tj. warianty wymowy, brzmienia. Jednak gdy na rynku pojawiły się ogólnodostępne telefony komórkowe, nazwa danego referentu i rozpowrzechnienie tego narzędzia komunikacji wahała się w granicach rzędu synonimicznego w nie mniej labylnym spektrum stylistycznym z mocnym przechyleniem w kierunku stylu potocznego:

сотый, сотик, сотовый, сотовый телефон, труба, трубка, мобила, телефон, мобильный, мобильник, мобильный телефон. Inni badacze wymieniają także *соточка, сотовик* (Мандрикова, Кротова 2013: 37). Z danego rzędu jedynie *сотовый, сотовый телефон, мобильный телефон* odnalazły się w przedziale stylu neutralnego z późniejszą przewagą użycia *мобильный телефон*. Wyraz *мобильник* przekroczył granice stylu potocznego i może być uważany za wyraz neutralny obok współczesnego *смартфон*. Repozytorium nie oferuje jednak możliwości obserwacji tematycznej słów. Wprowadzenie parametryki tematycznej możliwe jest w jego następnych edycjach RAS. Obecne parametry są wyszczegółowione w kolumnach 1.–9. tabeli 4.

XI. Struktura RAS oraz dostęp do zasobu

Dostęp do RAS jest zdalny – poprzez stronę internetową <https://jefed-repository.web.amu.edu.pl/> (nazwa użytkownika: *user*; hasło: *user*).

Poruszanie się po RAS jest możliwe w pionie i poziomie. Dane uzyskiwane w ten sposób umożliwiają odczyt całego artykułu hasłowego (w poziomie), jak całych zbiorów słów z określonym formantem (w pionie), przy czym poruszanie się możliwe jest po całym słowniku od tego samego wyrazu w indeksie a tergo do tego samego wyrazu w indeksie a fronte.

Tabela 4. Kolumny RAS w wymiarze pionowym (kolumny 1., 2.) i poziomym (3.–9.)

Wymiar pionowy RAS		Wymiar poziomy RAS						
1	2	3	4	5	6	7	8	9
Index a tergo: lemma without quotation mark	Index a fronte: lemma without quotation mark	Index a fronte: original lemma with quotation mark	Context with word form	SOURCE	YEAR	RFS	CITY	MAP

W dalszej kolejności będę omawiał każdą kolumnę z osobna. Przyjrzyjmy więc się poszczególnym poziomom oraz parametrom konsultacji w RAS.

XI.1. Wymiar pionowy RAS

Spójrzmy na tabelę 5. Wymiar pionowy – to wymiar poruszania się między indeksami w kolumnie 1. oraz 2. Indeks w kolumnie 1. – to lista wyrazów w układzie *a fronte* IAS, indeks w kolumnie 2. – to te same AS w innym układzie – *a tergo*, tj. w zwyczajnym układzie alfabetycznym umożliwiającym szukanie słów rozpoczynających się od określonej litery lub ich sekwencji w porządku alfabetycznym od lewej do prawej strony. Decyzja o umieszczeniu indeksu *a tergo* jako pierwszego z rzędu kolumn powstała podczas serii publicznych prezentacji RAS, w których uczestniczyło wielu lingwistów specjalizujących się w zakresie morfologii i słowotwórstwa. Lista ta umożliwia wyszukiwanie słów od końca wyrazów (tj. od prawej do lewej strony), co jest szczególnie ważne dla osób wyszukujących grupy wyrazów o określonych właściwościach gramatycznych, o określonych formantach. Postaci graficzne obu indeksów są tożsame w odróżnieniu od indeksu w trzeciej kolumnie, w której pisownia nie jest pozbawiona znaków cudzysłowu, uniemożliwiających alfabetyczne uporządkowanie danych.

Kolumna 1. Index *a tergo*: lemma without quotation mark (initial and terminal)

W danej kolumnie umieszczony jest IAS w układzie inwersyjnym *a tergo* w trzech porządkach: 1) numerycznym od 1 do 9, 2) alfabetycznym łacińskim od *A(a)* do *Z(z)* oraz alfabetycznym cyrylicznym od *A(a)* до *Я(я)* w kierunku od lewej do prawej strony ciągu znaków (wyrazu). Wyrazem takim jest lemma pozbawiona znaków cudzysłowu parnego (inicjalnego i medialnego) z wyjątkiem znaku cudzysłowu umieszczonego w środku wyrazu (medialnego). Wyrównanie w formatowaniu indeksu jest od strony prawej. Dany zabieg techniczny umożliwia oraz ułatwia wyszukiwanie słów od końca wyrazu zwłaszcza przy próbie wytypowania słów o tym samym formancie terminalnym. Każdy wyraz z danego indeksu jest połączony linkiem z tym samym wyrazem w kolejnej kolumnie, w której jest umieszczony indeks *a fronte* (por. niżej).

Kolumna 2. Index a fronte: lemma without quotation mark (initial and terminal)

Wyrazy w danym indeksie technicznie powiązane są z wyrazami w indeksie a tergo w kolumnie 1. Każdy AS w kolumnie 2. (*a fronte*) jest tożsamy z wyrazem w kolumnie 3. Różnica między AS w drugiej a trzeciej kolumnie a fronte polega na brakujących znakach cudzośćłowu.

Tabela 5. Indeksy a tergo i a fronte z obustronnymi hiperłączami – kolumny 1. oraz 2.

1	2
Index a fronte: lemma without quotation mark	Index a fronte: lemma without quotation mark
(...)	(...)
<u>лонгплей</u>	лонглист
лонг-плей	лонг-лист
лонг-листер	лонг-листер
лонглист	<u>лонгплей</u>
лонг-лист	лонг-плей
(...)	(...)

Podkreślę, że pierwsze dwa indeksy (kolumny 1. oraz kolumna 2.) to te same wyrazy w innym układzie alfabetycznym; są one skorelowane ze sobą systemem dwukierunkowych hiperłącz. Oznacza to, że użytkownik RAS, klikając w dowolny wyraz w indeksie a tergo, zostanie przekierowany na ten sam wyraz w kolumnie drugiej – w indeksie a fronte. I odwrotnie. Pierwsze dwa indeksy w pierwszych dwóch kolumnach (kolumny 1. oraz 2.) są robocze, ułatwiają pionową nawigację po RAS – por. tabelę 6.

XI.2. Wymiar poziomy RAS

Wymiar poziomy liczy siedem kolumn odgradzonych od pierwszych dwóch roboczych indeksów wypukłą pionową linią. Jeden wiersz w kolumnach 3.–7. prezentuje sobą pojedynczy artykuł hasłowy. Przyjrzyjmy się mu z bliska.

Kolumna 3. Index a fronte: original lemma with quotation mark

Jest to podstawowy i centralny pion indeksu z oryginalnym zapisem graficznym zmodyfikowanym do postaci lemmowej. Wyrazy w indeksie są opatrzone pogrubieniem, ale nie są opatrzone linkiem.

Tabela 6. Kolumna 3. w RAS

3		
Index a fronte: lemma without quotation mark	Index a fronte: lemma without quotation mark	Index a fronte: lemma with quotation mark
(...) лонгплей лонг-плей лонг-листер лонглист лонг-лист (...)	(...) лонглист лонг-лист лонг-листер лонгплей лонг-плей (...)	(...) лонглист "лонг-лист" лонг-листер лонгплей лонг-плей (...)

Po kliknięciu na wyrazy w kolumnach 1. Lub 2. pozostaje „śląd” w postaci zmienionego koloru czcionki z podkreśleniem. Ułatwia to pojedyncze korzystanie z zasobu.

Kolumna 4. Context with word form

Dane kontekstowe zawierają zdanie lub kilka zdań w oryginalnej pisowni, w której figuruje oryginalna słowoforma oraz postać graficzna AS wyróżniona pogrubieniem czcionki – por. poniższa tabela:

Tabela 7. Kolumna 4. zawierająca kontekst użycia AS

4
Context
Уже никто из военкоматчиков не пытается воздействовать на призывника высокими фразами о патриотизме, о священном долге перед Родиной и т.д.

Kolumna 5. Source

W danym poziomie znajduje się tytuł gazety regionalnej, z której został pobrany kontekst z AS. Większość tytułów ma również przydzielony oryginalny numer woluminu.

Kolumna 6. Year

Rok jest datacją ukazania się numeru gazety zawierającym kontekst użycia DA.

Kolumna 7. RFS

W danej kolumnie figuruje nazwa subiektu Federacji Rosyjskiej, skąd pochodzi gazeta regionalna.

Kolumna 8. City

Tytuł kolumny wskazuje na miasto ukazania się woluminu gazety regionalnej wskazanej w kolumnie 7.

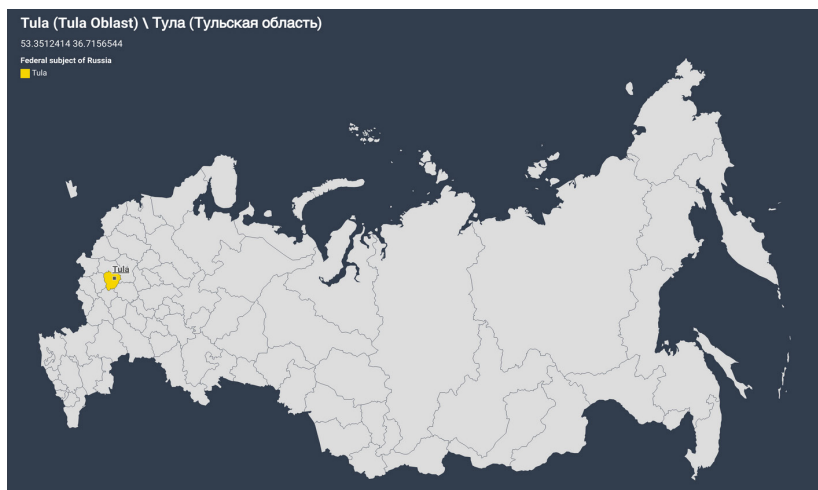
Kolumny 5., 6., 7., 8. zaprezentowane są dodatkowo w tabeli 8:

Tabela 8. Kolumny 5., 6., 7., 8. przedstawiające metadane: source (title of newspaper), year, RFS and name of the city in this RFS

5	6	7	8
Source	Year	RFS	City
<Пермские новости №31 (1203)	2003	Пермский край	Пермь

Kolumna 9. MAP

Ostatnia kolumna z mapą Federacji Rosyjskiej. Mapa jest ukryta. Ukazuje się czytelnikom po naprowadzeniu kursora na zapis [map](#). Mapa w formacie PNG ukazuje RFS (w kolorze żółtym), miasto w RFS (jako punkt – czarny kwadrat) oraz współrzędne w systemie dziesiętnym (szerokość i długość geograficzna), np. 56.6295264 40.5700972. Mapa jest utworzona w technologii designerskiej DATAWRAPPER (datawrapper.de) – por. poniższą infografikę:



Rysunek 14. Mapa miasta Tuła w obwodzie Tulskim [oprac. własne]

Zapraszam do korzystania z Repozytorium agnonimów słownikowych, jego pierwszej edycji. Wszystkie uwagi do danych RAS proszę kierować na adres jerfed@amu.edu.pl. Ponadto zachęcam Czytelników do podzielenia się pomysłami związanymi z rozbudową oraz wszelką optymalizacją RAS.

Bibliografia

W niniejszym rozdziale notacja obejmuje pozycje w kilku różnych językach, w tym – w kilku alfabetach. W związku z tym podjęto decyzję o umieszczeniu w *Literaturze naukowej* pozycji w alfabecie łacińskim z podaniem – w nawiasach kwadratowych – imion autorów oraz tytułów w alfabecie oryginalnym. W rozdziale *Słowniki weryfikacyjne* umieszczono jedynie oryginalne zapisy pozycji.

Literatura naukowa

- Alatortseva, S.I.** [Алаторцева, С.И.]. 1999. *Problemy neologii i russkaya neografiya. (Avtoreferat doktorskoy dissertatsii).* [Проблемы неологии и русская неография (Автореферат докторской диссертации)]. Sankt Peterburg: LGOU im. A.S. Pushkina.
- Aleksandrzak, M., red.** 2018. *Języki specjalistyczne w lingwistyce stosowanej: między teorią i praktyką.* Seria wydawnicza „Język – Kultura – Komunikacja”, Vol. 23. Poznań: Wydawnictwo Naukowe UAM.
- Alekseyev, V.A.** [Алексеев, В.А.]. 1973. *Voprosy stilistiki. Problemy zhurnalistiki* 2 (11), 41–49.
- Babenko, L.G., red.** [Бабенко, Л.Г.]. 2005. *Bol'shoy tolkovyy slovar' russkikh sushchestvitel'nykh. Ideograficheskoye opisaniye. Sinonimy, antonimy.* [Большой толковый словарь русских существительных. Идеографическое описание. Синонимы, антонимы.]. Moskva: AST-PRESS KNIGA.
- Bańko, M., Hebal-Jeziarska M.** 2015. *Wariancja graficzna w procesie adaptacji zapożyczeń.* *Bulletin de la Société Polonaise de Linguistique* XXI, 141–152.
- Bartkov, V.I., Minina, L.I.** [Бартков, В.И., Минина, Л.И.]. 2019. *Derivatariy (kolichestvennyy "slovar" affiksov) sovremennogo russkogo yazyka: v 2 t. T. 1: Trista prepozitivnykh derivatsionnykh morfem.* [Дериватарий (количественный "словарь" аффиксов) современного русского языка: в 2 т. Т. 1: Триста препозитивных деривационных морфем]. Kazan: Buk.
- Bartmiński, J.** 1980. *Słownik ludowych stereotypów językowych. Zeszyt próbny.* Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego.
- Bartwicka, H., Fedorushkov, Y., Małek, E., Wawrzyńczyk, J.** 2007. *Slovarnyye agnonimy russkogo yazyka.* [Словарные агнонимы русского

- языка]. *Seria: Semiosis Lexicographica* XLIII. Warszawa: Wydawnictwo Takt.
- Bieleń, S., red. 1997.** *Patrząc na Wschód z problematyki polityki wschodniej III RP*. Warszawa: Centrum Badań Wschodnich UW.
- Bierwiazzonek, B. 2006.** *O języku ucieleśnionym*. W: O. Sokołowska, D. Stanilewicz, red. *Językoznawstwo kognitywne III. Kognitywizm w świetle innych teorii*. Gdańsk: Wydawnictwo Uniwersytetu Gdańskiego, 444–479.
- Czubiński, A. 1997.** *Europa XX wieku*. Poznań: Wydawnictwo Poznańskie.
- Donaj, Ł. 2001.** *Media w systemie politycznym w Federacji Rosyjskiej*. Poznań: Wydawnictwo UAM.
- Drawicz, A. 1997.** Przewartościowanie w Rosji w polityce III RP. W: S. Bieleń, red. *Patrząc na Wschód z problematyki polityki wschodniej III RP*. Warszawa: Centrum Badań Wschodnich Uniwersytetu Warszawskiego, 48–62.
- Fedorushkov, Y. 2005.** Automatyczna analiza morfologiczna jako narzędzie oceny poprawności wyrazów języka rosyjskiego. *Investigationes Linguisticae*, Vol. XII, 33–43.
- Fedorushkov, Y. 2007.** *Methods for electronic excerption of new words in Russian*. W: B. Lewandowska-Tomaszczyk, red. *PALC 2007: Practical Applications in Language and Computers. Papers from the International Conference at the University of Łódź, 19-22 April 2007*. Frankfurt am Main: Peter Lang Verlag, 163–196.
- Fedorushkov, Y. 2008.** *Metody ekscerpcji neologizmów języka rosyjskiego (na materiale współczesnej prasy rosyjskiej)*. [Praca doktorska; promotor – P. Wierchoń], <http://hdl.handle.net/10593/16715> – WEB: 11.05.2023.
- Fedorushkov, Y. 2009.** Ekscerpcja wyrazów cudzojęzycznych z tekstów prasy rosyjskiej. *Seria: Semiosis Lexicographica*, Vol. LII, Warszawa: Wydawnictwo Takt.
- Fedorushkov, Y. 2009.** *Słowniki elektroniczne oraz środowisko hipertekstowe w leksykografii a Słownik języka polskiego pod red. W. Doroszewskiego*. W: J. Wawrzyńczyk, red. *Czterdzieści lat minęło nad „Słownikiem Doroszewskiego”*, Warszawa: Wydawnictwo Takt, 27–37.
- Fedorushkov, Y. 2010.** Metody automatyzacji ekscerpcji konstrukcji atrybutywnych języka rosyjskiego. W: W. Chlebda, red. *Na tropach reproduktów. W poszukiwaniu wielowrazowych jednostek języka*, Opole: Wydawnictwo Uniwersytetu Opolskiego, 59–85.
- Fedorushkov, Y. 2019.** *Adaptacja warsztatu leksykograficznego do automatyzacji rozpoznawania frazemów w systemie uczącym się. Na przykładzie konstrukcji werbo-nominalnych języka rosyjskiego*. Poznań: Wydawnictwo Naukowe UAM.

- Fedorushkov, Y. 2018a.** *Kieszonkowy słownik piłki nożnej rosyjsko-polski i polsko-rosyjski*. Warszawa: Centrum Polsko-Rosyjskiego Dialogu i Porozumienia.
- Fedorushkov, Y., 2018b.** Prolegomena do tagowania frazemów w równoległym korpusie rosyjsko-polskim (literatura piękna) w aspekcie przekładoznawczym. *Acta Polono-Ruthenica*, Vol. 2/XXIII, 55–74.
- Fedorushkov, Y., 2018c.** Zwroty czasownikowo-rzeczownikowe o tematyce futbolowej w słowniku rosyjsko-polskim /polsko-rosyjskim. *Studia Rossica Posnaniensia*, Vol. XLIII, 45–62.
- Fedorushkov, Y., Dzienisiewicz, D. 2014.** Automatyizacja wizualizacji grafowej synonimów dla potrzeb dydaktyki języków obcych (na przykładzie przymiotników rosyjskich z prefiksem без-/без-). *Kultury Wschodniosłowiańskie – Oblicza i Dialog*, Vol. 4. 35–48.
- Fedorushkov, Y., Dzienisiewicz, D. 2019.** Metoda wyodrębniania odtwarzalnych konstrukcji werbo-nominalnych języka rosyjskiego z korpusów tekstowych. *Studia Rossica Posnaniensia*, Vol. 44(2), 379–394.
- Fedorushkov, Y., Graliński, F., Jurkiewicz, D. 2018.** *Prolegomena do automatyzacji rozpoznawania rosyjskich frazemów werbo-nominalnych*. W: M. Aleksandrak, red. 2018. Języki specjalistyczne w lingwistyce stosowanej: między teorią i praktyką. Seria wydawnicza „Język – Kultura – Komunikacja”, Vol. 23. Poznań: Wydawnictwo Naukowe UAM, 93–129.
- Fedorushkov, Y., Narloch, A. 2014.** Prolegomena do dydaktycznej prezentacji konceptu językowego w wizualizacji grafowej (na przykładzie rosyjskiego konceptu BELYJ), *Studia Rossica Gedanensia*, Vol. 1, 179–208.
- Fedorushkov, Y., Szutkowski, T.** [Федорушков, Ю., Шутковский, Т.]. **2010.** Leksiko-grammatyckaja sochetaemost' atributivnyh slovosochetaniy russkogo jazyka v kontekste metodov komp'juternoj ekscerptcii. [Лексико-грамматическая сочетаемость атрибутивных словосочетаний русского языка в контексте методов компьютерной эксцерпции.]. W: T.R. Guzman, L. Sokolova, I. Voytakova, red., II. Mezhdunarodnaya konferentsiya "Russkij yazyk i literatura v mezhdunarodnom obrazovatel'nom prostranstve: sovremennoe sostoyanie i perspektivy", Granada, 8–10 sentyabrya 2010 g., Vol. II. Granada: Rubiños – 1860, S.A, 1564–1569.
- Gabdreeva, N.V., Gurčiani, M.T.** [Габдреева, Н.В., Гурчиани, М.Т.]. **2012.** *Slovar' kompozitov russkogo yazyka novejšego perioda*. [словарь комpositов русского языка новейшего периода.]. Moskva: FLINTA / Nauka.
- Handke, K. 1976.** Budowa morfologiczna i funkcje compositów polskich (z uwzględnieniem innych języków zachodniosłowiańskich). *Prace Językoznawcze*, Vol. 81. Wrocław – Zakład Narodowy im. Ossolińskich.

- Hluszko, A. 2016.** Wstrząsająca treść jako manipulacyjny element dyskursu konfliktu. *Oblicza komunikacji*, Vol. 9, 135–143.
- Hrynkiewicz-Admaskich, B. 2005.** *Rosyjsko-polski słownik innowacji leksykalnych*. Poznań: Wydawnictwo Naukowe UAM.
- Jodłowski, S. 1936.** Z zagadnień związanych z nauczaniem gramatyki polskiej. Wyrazy z. przedrostkami oraz. tzw. zrosty i złożenia. *Język Polski*, Vol. XVI, 103–110.
- Kalیشان, J. [Калишан, Е.]. 1980.** *Internatsional'nyye prepozitivnyye morfe-mygreko-latinskogo proiskhozhdeniya v sovremennoy russkom i pol'skom slovoobrazovanii*. [*Интернациональные препозитивные морфемы греко-латинского происхождения в современном русском и польском словообразовании*]. Poznań: Wydawnictwo Naukowe UAM.
- Kobozeva, I.M. [Кобозева, И.М.]. 2000.** *Leksicheskaya semantika*. [*Лексическая семантика*]. Moskva: Editorial URSS.
- Kostomarov, V.G. [Костомаров, В.Г.]. 1971.** *Russkiy yazyk na gazetnoy polose. Nekotoryye osobennosti yazyka sovremennoy gazetnoy publitsistiki*. [*Русский язык на газетной полосе. Некоторые особенности языка современной газетной публицистики*]. Moskva: Izdatel'stvo Moskovskogo universiteta.
- Kostomarov, V.G. [Костомаров, В.Г.]. 1998.** *Âzykovoj vkus èpohi*. [*Языковой вкус эпохи*]. *Russkiy yazyk v shkole*, Vol. 1, 11–13.
- Kostomarov, V.G. [Костомаров, В.Г.]. 1999.** *Âzykovoj vkus èpohi. Iz nablyudeniya nad rechevoy praktikoy mass-media*. [*Языковой вкус эпохи. Из наблюдений над речевой практикой масс-медиа*]. Sankt-Peterburg: Zlatoust.
- Kotelova, N.Z. [Котелова, Н.З.]. 2015.** *Izbrannyye raboty*. [*Избранные работы*]. Sankt-Peterburg: Izdatel'stvo Nestor-Istoruya.
- Kotelova, N.Z. [Котелова, Н.З.]. red. 1995.** *Slovar' novykh slov russkogo âzyka (seredina 50-h – seredina 80-h godov)*. [*Словарь новых слов русского языка (середина 50-х – середина 80-х годов)*]. Sankt-Peterburg: Dmitriy Bulanin.
- Kozhina, M.N. [Кожина, М.Н.]. 2003.** *Rehevaya sistemnost' funktsional'nogo stilya*. W: Kozhina, M.N. red., 2003, 5–20.
- Kozhina, M.N., [Кожина, М.Н.] red. 2003.** *Stilisticheskyy entsiklopedicheskyy slovar' sovremennoy russkogo yazyka*. Moskva: Flinta, Nauka.
- Krasil'shik I.S., Rahilina E.V. [Красильщик, С.И., Рахилина, Е.В.]. 1992.** *Predmetnyye imena v sisteme "Leksikograf"*. [*Предметные имена в системе „Лексикограф“*]. *NTI*, Vol. 9, iss. 2, 24–31.
- Kustova, G., Lyashevskaya, O. Paducheva, Ye., Rakhilina, Ye., Shemayeva, O. [Кустова, Г., Ляшевская, О. Падучева, Е., Рахилина, Е.,**

- Шеманаева, О.]. 2005. *Semanticheskaya razemtka leksiki v Natsional'nom korpusе russkogo yazyka: printsipy, problemy, perspektiv*. [Семантическая раземтка лексики в Национальном корпусе русского языка: принципы, проблемы, перспективы.]. W: V.A. Plungyan, red. 2005. *Национальный корпус русского языка: 2003–2005*. Результаты и перспективы. Moskva: Indrik.
- Kustova, G., I. Paducheva, Ye.V.** [Кустова, Г.И., Падучева, Е.В.]. **1994**. Slovar' kak leksičeskaâ baza dannyh. [Словарь как лексическая база данных.]. *Voprosy âzykoznanîâ*, Vol. 4, 96–106.
- Malcówna, M.** **1974**. Dwuczłonowe przezwiska i przydomki staropolskiej. *Onomastica*, Vol. XIX, 191–227.
- Mandrikova, G.M., Krotova, A.G.** [Мандрикова, Г.М., Кротова, А.Г.]. **2013**. *Sovremennyy russkiy yazyk: leksikologiya vskhemakh, zadaniyakh i uprazhneniyakh: uchebno-metodicheskoye posobiye*. [Современный русский язык: лексикология в схемах, заданиях и упражнениях: учебно-методическое пособие.]. Novosibirsk: Izdatel'stvo NGTU.
- Marcińczuk, M., Kocoń, J. Ptak, M., Kaczmarek, A., 2010**. *Inforex*, CLARIN-PL digital repository, <http://hdl.handle.net/11321/13> – WEB: 12.12.2022 (web resource – CLARIN-PL digital repository).
- Marcińczuk, M., Oleksy, M., Kocoń, J.** **2017**. *Inforex – a Collaborative System for Text Corpora Annotation and Analysis*. W: R. Mitkov, G. Angelova, red. *Proceedings of the International Conference Recent Advances in Natural Language Processing, RANLP 2017*. Varna: INCOMA Ltd, 473–482.
- Maziarz M., Piasecki M., Szpakowicz S., Rabiega-Wiśniewska J.** **2011**. Semantic Relations Among Nouns in Polish Wordnet Grounded in Lexicographic and Semantic Tradition. *Cognitive Studies/Études Cognitives*, Vol. 11, 161–181.
- Mietła, J.** **2001**. Zjawisko hybrydyzacji w czeskiej literaturze językoznawczej. *Bohemistyka*, Vol. 1, 52–60.
- Misturska-Bojanowska, J.** [Мистурса-Бояновска, Й.]. **2013**. *Dinamiche-skiye protsessy v slovoobrazovanii imensushchestvitel'nykh v russkom yazyke 2-oy pol. XX veka*. [Динамические процессы в словообразовании имен существительных в русском языке 2-ой пол. XX века.]. Szczecin: Wydawnictwo Naukowe Uniwersytetu Szczecińskiego.
- Morkovkin, V.V., Morkovkina, A.V.** [Морковкин, В.В., Морковкина, А.В.]. **1997**. *Russkie agnonimy: Slova, kotorye ty ne znaet*. [Русские агнонимы. Слова, которые мы не знаем]. Moskva: Russkiy yazyk.
- Nemchenko, V.N.** [Немченко, В.Н.]. **1984**. *Sovremennyy russkiy yazyk. Slovoizmeneniye*. [Современный русский язык. Словообразованиё.]. Moskva: Vysshaya shkola.

- Novikova, O.V.** [Новикова, О.В.]. **2002.** *Strukturno-funktsional'naya kharakteristika russkikhneologizmov dvukh stilisticheskikh podsistem kontsa XX – nachala XXI vv. (na materiale gazetnogo diskursa i komp'yuternogo slenga).* [Структурно-функциональная характеристика русских неологизмов двух стилистических подсистем конца XX – начала XXI вв. (на материале газетного дискурса и компьютерного сленга)]. [thesis, supervisor – O. Lagut.]. Novosibirsk: Novosibirskij Gosudarstvennyj universitet, http://zipsites.ru/books/rus_neologizmy – WEB: 13.12.2005.
- Panov, M.V.** [Панов, М.В.]. **1988.** *Iz nablyudeniy nad stilemsegodnyashney periodiki. Yazyk sovremennoy publitsistiki.* [Из наблюдений над стилем сегодняшней периодики. Язык современной публицистики.]. Moskva: Nauka.
- Plungyan, V.A.** [Плунгян, В.А.]. **red. 2005.** *Natsional'nyu korpus russkogo yazyka: 2003–2005. Rezul'taty i perspektivy.* [Национальный корпус русского языка: 2003–2005. Результаты и перспективы.]. Moskva: Indrik..
- Plungyan, V.A.** [Плунгян, В.А.]. **red. 2009.** *Natsional'nyu korpus russkogo yazyka: 2006–2008. Rezul'taty i perspektivy.* [Национальный корпус русского языка: 2006–2008. Новые результаты и перспективы.]. Sankt-Petersburg: Nestor-Istoriya.
- Popov, R.N.** [Попов., Р.Н.]. **1996.** *Novyye slova i slovosochetaniya v yazyke sovremennoy pressy.* [Новые слова и словосочетания в языке современной прессы.]. *Russkiy yazyk v shkole*, Vol. 1, 15–16.
- Popova, T.V.** [Попова., Р.Н.]. **2005.** *Russkaya neologiya i neografiya.* [Русская неология и неография]. Ekaterinburg: UGTU–UPI, http://window.edu.ru/window_catalog/files/r28514/ustu121.pdf – WEB: 13 06 2005.
- Popova, T.V., Raciburskaâ L.V., Gugunava, D.V.** [Попова, Т.В., Рацибурская Л.В., Гугунава, Д.В.]. **2005.** *Neologiya i neografiya sovremennogo russkogo yazyka. Uchebnoye posobiye.* [Неология и неография современного русского языка. Учебное пособие.]. Moskva: Flinta, Nauka.
- Pospiszyl, A.** **2000.** *Composita rzeczownikowe w słowniku gwarowym Andrzeja Cinciały.* *Prace Językoznawcze*, Vol. 2, 139–148.
- Puzynina, J.** **1968.** *Co to jest onomazjologia?* *Lingua Poznaniensis*, Vol. XII/ XIII, 119–124.
- Rakhilina, Ye.V., Kustova, G.I., Lyashevskaya, O.N., Reznikova, T.I., Shemanayeva, O.Yu.** [Рахилина, Е.В., Кустова, Г.И., Ляшевская, О.Н., Резникова, Т.И., Шеманаева, О.Ю.]. **2009.** *Zadaçi i principy semantičeskoj razmetki leksiki v NKRÁ.* [Задачи и принципы разметки лексики в НКРЯ.]. W: V. A. Plungân, red. *Nacional'nyj korpus russkogo âzyka:*

- 2006–2008. Novye rezul'taty i perspektivy. Sankt-Petersburg: Nestor, Istoriya., 215–239.
- Rahmanova, L.I., Suzdal'ceva, V.N.** [Рахманова, Л.И., Суздальцева, В.Н.]. 1997. *Sovremennyy russkiyazyk. Leksika. Frazeologiya. Morfologiya: Uchebnoye posobiye*. [Современный русский язык. Лексика. Фразеология. Морфология: Учебное пособие.]. Moskva: Izdatel'stvo Moskovskogo Gosudarstvennogo Universiteta.
- Rozental', D.Ė.** [Розенталь, Д.Е.]. 1977. *Praktičeskaâ stilistika russkogo âzyka*. [Практическая стилистика русского языка.]. Moskva: Vysshaya shkola.
- Rudnik-Karwatowa, Z.** 2003. O najnowszych tendencjach w słowotwórstwie, Procesy innowacyjne w językach słowiańskich, *Prace Slawistyczne*, Vol. 114, 199-210.
- Sinyachkina, L.** [Синячкина, Л.]. 1997. *Osnovnye tendencii v razvitii leksiki âzyka gazety na sovremennom ètape (1985–1996 gg.)*. Avtoreferat dissertacii. [Основные тенденции в развитии лексики языка газеты на современном этапе, 1985–1996 гг. Автореферат диссертации.]. Moskva: Moskovskij Gosudarstvennyj Universitet.
- Sklârevskaâ, G.N.** [Сляревская, Г.Н.]. red., 2005. *Tolkovyj slovar' russkogo âzyka načala XXI veka. Aktual'naâ leksika: ok. 8500 slov i ustojčivyh slovosocetaniĭ*. [Толковый словарь русского языка начала XXI века. Актуальная лексика: около 8500 слов и устойчивых словосочетаний.]. Sankt-Peterburg: Ėksmo.
- Sosnowski, W.** 2011. *Tendencje analityczne we współczesnym języku polskim i rosyjskim*. [Praca doktorska; promotor – Violetta Koseska-Toszewa]. Warszawa: POLSKA Akademia Nauk/ Instytut Slawistyki – WEB: <https://hdl.handle.net/20.500.12528/42> – WEB: 21 02 2023.
- Trávníček, F.** 1948. *Mluvnice spisovné češtiny*, T I. Praha: Slovanské nakladatelství.
- Trofimova, G.N.** [Трофимова, Г.Н.]. 2001a. Как обозначит' принадлежность к Internetu. [Как обозначить принадлежность к Интернету.]. *Mir russkogo slova*, Vol. 4, 73–75.
- Trofimova, G.N.** 2001b. *K voprosu o spetsifike funkcionirovaniya russkogo yazyka v Internete (norma i uzus)*. [К вопросу о специфике функционирования русского языка в Интернете (норма и узус)]. W: Trudy Mezhdunarodnogo seminarâ Dialog-2001 po komp'yuternym tekhnologiyam. <http://www.dialog-21.ru> – WEB: 06.12.2009.
- Vepreva, I.T.** [Вепрева, И.Т.]. 2005. *Yazykovaya refleksiya v post-sovet'skiyu epokhu*. [Языковая рефлексия в постсоветскую эпоху.]. Moskva: OLMA-PRESS.

- Waszakowa, K. 2005.** *Przejawy internacjonalizacji w słowotwórstwie współczesnej polszczyzny*. Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.
- Waszakowa, K. 2015.** *Kognitywistyczne ujęcia derywatów słowotwórczych a koncepcja Miloša Dokulila*. *Biuletyn Polskiego Towarzystwa Językoznawczego*, Vol. LXXI, 111–126.
- Waszakowa, K. 2021.** *Język w działaniu i działania na języku: szkice semantyczno-słowotwórcze*. 1st red. Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.
- Wawrzyńczyk, J., Fedorushkov, Y., Ignasiak, A., Kulpina, W., Rodak, A., Siemanowska, U., Tatarinów, W.** *Nazvaniya zhenshchin. Dopolnitel'nyy slovar' nyu slovar'*. [*Названия женщин. Дополнительный словарь*]. Seria: *Semiosis Lexicographica*, Vol. XXXV. Warszawa: Wydawnictwo Takt, 2006.
- Wieruchoń, P. 2008.** *Fotodokumentacja. Chronologizacja. Emendacja. Teoria i praktyka weryfikacji materiału leksykalnego w badaniach lingwistycznych*. Poznań: SORUS.
- Wieruchoń, P. 2009.** *Fotodokumentacja 3.0*. W: P. Nowak, P. Nowakowski, red. *Język, Komunikacja, Informacja*, Vol. 4, 63–80.
- Wieruchoń, P., 2010.** *Torując drogę teorii lingwochronologizacji, Investigations Linguisticae*, Vol. XX, 105–185.
- Wieruchoń, P., red., 2010.** *Depozytorium leksykalne języka polskiego. Nowe fotomateriały z lat 1901–2010 – Tom 1*. Warszawa: BEL Studio.
- Wieruchoń, P., Graliński F. 2016.** *Z kart historii „parcia na” neologizmy, Poradnik Językowy*, Vol. 4, 110–129.
- Zaliznyak, A.A. [Зализняк, А.А.]. 1987.** *Grammaticheskiĭ slovar' russkogoazyka. Slovoizmeneniye: okolo 100 000 slov*. [*Грамматический словарь русского языка. Словоизменение: около 100 000 слов*]. Moskva: Russkiy yazyk.
- Zaliznyak, A.A. [Зализняк, А.А.]. 2022.** *Grammaticheskiĭ slovar' russkogoazyka. Slovoizmeneniye: okolo 110000 slov*. [*Грамматический словарь русского языка. Словоизменение: около 110 000 слов*]. VI th red. with on-line version. Moskva: AST-Press, <https://gramdict.ru/> – WEB: 06.12.2022.

Słowniki weryfikacyjne

БАС (большой академический словарь) –

Чернышёв, В.И. et al. 1950–1965. *Словарь современного русского литературного языка*. Т. 1 (А – Б) – Т. 15 (Х – Я). Москва / Ленинград: Издательство Академии Наук СССР.

БАС – kontynuacja:

- Чернышёв, В.И.** 1951. *Словарь современного русского литературного языка* (В).
- Чернышёв, В.И.** 1954. *Словарь современного русского литературного языка* (Г – Е).
- Чернышёв, В.И.** 1955. *Словарь современного русского литературного языка* (Ж – З).
- Чернышёв, В.И.** 1956. *Словарь современного русского литературного языка* (И – К).
- Чернышёв, В.И.** 1957. *Словарь современного русского литературного языка* (Л – М).
- Чернышёв, В.И.** 1958. *Словарь современного русского литературного языка* (Н).
- Чернышёв, В.И.** 1959. *Словарь современного русского литературного языка* (О).
- Чернышёв, В.И.** 1959. *Словарь современного русского литературного языка* (П – пнуть).
- Чернышёв, В.И.** 1960. *Словарь современного русского литературного языка* (По – поясок).
- Чернышёв, В.И.** 1961. *Словарь современного русского литературного языка* (Пра – пятью).
- Чернышёв, В.И.** 1961. *Словарь современного русского литературного языка* (Р).
- Чернышёв, В.И.** 1962. *Словарь современного русского литературного языка* (С – сняться).
- Чернышёв, В.И.** 1963. *Словарь современного русского литературного языка* (Со – сям).
- Чернышёв, В.И.** 1963. *Словарь современного русского литературного языка* (Т).
- Чернышёв, В.И.** 1964. *Словарь современного русского литературного языка* (У – Ф).
- Ковтун, Л.С., Петушков, В.П.** 1965. *Словарь современного русского литературного языка* (Х – Я).
- Брагина, А.А.** 1997. *Новое в русской лексике. Словарные материалы – 1988. Филологические науки* 2, 111–115.
- Буцева, Н.В. ed.** 2004. *Новое в русской лексике. Словарные материалы – 1990.* Санкт-Петербург: Дмитрий Буланин.
- Плотицын, В., Соловьев, Н., Юдина, А., Судоплатова, М., Бояркина, В., Козулина, Н.** 2001. *Новое в русской лексике. Словарные материалы – 1989.* Санкт-Петербург: Дмитрий Буланин.

- Денисенко, Ю.Ф. ed. 2005.** *Новое в русской лексике. Словарные материалы – 1991.* Санкт-Петербург: Дмитрий Буланин.
- Денисенко, Ю.Ф. ed. 2006.** *Новое в русской лексике. Словарные материалы – 1994.* Санкт-Петербург: Дмитрий Буланин.
- Евгеньева, А.П. ed. 2001.** *Словарь синонимов русского языка. Т. I–II.* Москва: Астрель.
- Зализняк, А.А. 1987.** *Грамматический словарь русского языка. Словоизменение: около 100000 слов.* Wyd. III zmodyfikowane i uzupełnione. Москва: Русский язык..
- Зализняк, А.А. 2022.** *Грамматический словарь русского языка. Словоизменение: около 110000 слов.* Wyd. VI stereotypowe względem wyd. IV. Москва: АСТ-Пресс. Oficjalna wersja internetowa słownika: <https://gramdict.ru/> – WEB: 06.12.2022.
- Котелова, Н.З. ed. 1980.** *Новое в русской лексике. Словарные материалы – 1977.* Москва: Русский язык.
- Котелова, Н.З. ed. 1981.** *Новое в русской лексике. Словарные материалы – 1978.* Москва: Русский язык.
- Котелова, Н.З. ed. 1982.** *Новое в русской лексике. Словарные материалы – 1979.* Москва: Русский язык.
- Котелова, Н.З. ed. 1984a.** *Новое в русской лексике. Словарные материалы – 1980.* Москва: Русский язык.
- Котелова, Н.З. ed. 1984b.** *Новые слова и значения. Словарь-справочник по материалам прессы и литературы 70-х гг.* Москва: Русский язык.
- Котелова, Н.З. ed. 1986a.** *Новое в русской лексике. Словарные материалы – 1981.* Москва: Русский язык.
- Котелова, Н.З. ed. 1986b.** *Новое в русской лексике. Словарные материалы – 1982.* Москва: Русский язык.
- Котелова, Н.З. ed. 1987.** *Новое в русской лексике. Словарные материалы – 1983.* Москва: Русский язык.
- Котелова, Н.З. ed. 1989.** *Новое в русской лексике. Словарные материалы – 1984.* Москва: Русский язык.
- Котелова, Н.З. ed. 1995.** *Словарь новых слов русского языка (середина 50-х–середина 80-х гг.).* Санкт-Петербург: Дмитрий Буланин.
- Котелова, Н.З. ed. 1996.** *Новое в русской лексике. Словарные материалы–1986.* Санкт-Петербург: Дмитрий Буланин.
- Котелова, Н.З., Денисенко, Ю.Ф. eds. 1996a.** *Новое в русской лексике. Словарные материалы–1985.* Москва: Дмитрий Буланин.
- Котелова, Н.З., Денисенко, Ю.Ф. eds. 1996b.** *Новое в русской лексике. Словарные материалы–1987.* Москва: Дмитрий Буланин.

- Котелова, Н.З., Сорокин, Ю.С. ed. 1971.** *Новые слова и значения. Словарь-справочник по материалам прессы и литературы 60-х гг.* Москва: Советская Энциклопедия.
- Левашов, Е.А. ed. 1996a.** *Новое в русской лексике. Словарные материалы – 1988.* Санкт-Петербург: Дмитрий Буланин.
- Левашов, Е.А. ed. 1996b.** *Новое в русской лексике. Словарные материалы – 1988.* Москва: Дмитрий Буланин.
- Левашов, Е.А. ed. 1997.** *Новые слова и значения. Словарь-справочник по материалам прессы и литературы 80-х гг.* Санкт-Петербург: Дмитрий Буланин.
- Левашов, Е.А. ed. 2004.** *Новое в русской лексике. Словарные материалы – 1992.* Санкт-Петербург: Дмитрий Буланин.
- МАС (малый академический словарь) –**
- Евгеньева, А.П. ed. 1981.** *Словарь русского языка. В четырёх томах. Т. I (А – Й).* Wyd. II. Москва: Русский язык.
- Евгеньева, А.П. ed. 1982.** *Словарь русского языка. В четырёх томах. Т. II (К – О).* Wyd. II. Москва: Русский язык.
- Евгеньева, А.П. ed. 1983.** *Словарь русского языка. В четырёх томах. Т. III (П – Р).* Wyd. II. Москва: Русский язык.
- Евгеньева, А.П. ed. 1984.** *Словарь русского языка. В четырёх томах. Т. IV (С – Я).* Wyd. II. Москва: Русский язык.
- Милейковская, Г., Казем-Бек, Г., Домбровская И., Гаевская, Л. 1988.** *Материалы для словаря русского языка 70-х и 80-х годов XX века не зарегистрированные в нормативных словарях. Т. 1.* Warszawa: Wydawnictwo Uniwersytetu Warszawskiego.
- Милейковская, Г., Петельщиц, Р., Тарновская, И., Тимошук, М. 1996.** *Материалы для словаря русского языка 80-х годов XX века не зарегистрированные в нормативных словарях. Т. 3.* Warszawa: Energeia.
- Милейковская, Г., Петельщиц, Р., Тимошук, М. 1990.** *Материалы для словаря русского языка 70-х и 80-х годов XX века не зарегистрированные в нормативных словарях. Т. 2.* Warszawa: Wydawnictwo Uniwersytetu Warszawskiego.
- Ожегов, С.И., Шведова, Н.Ю. 1992.** *Толковый словарь русского языка.* Москва: Русский язык.
- Ушаков, Д.Н. red., 1996.** *Толковый словарь русского языка.* Москва: ТЕРРА.
- Ушаков, Д.Н., Крючков, С.Е. 1980.** *Орфографический словарь.* Москва: Просвещение.

Korpusy i inne zasoby tekstowe

- Piasecki, M. 2018.** *WordnetLoom 2.0* – WEB: <http://hdl.handle.net/11321/557> – WEB: 12.12.2022 (web resource – CLARIN-PL digital repository).
- Pawłowski, A.T. 2016.** *ChronoPress – Chronologica Corpus*, <http://hdl.handle.net/11321/260> – WEB: 12.12.2022 (web resource – CLARIN-PL digital repository).
- Pawłowski, A. 2021.** *ChronoPress – Korpus Tekstów Prasowych*, doi:10.34616/139101. <http://chronopress.clarin-pl.eu/> – WEB: 12.12.2022 (web resource – CLARIN-PL digital repository).
- Fedorushkov, Y. 2023.** *Corpus of Russian Local Press of the Millennium Period (1996–2006)*, CLARIN-PL digital repository – WEB: <http://hdl.handle.net/11321/911> – WEB: 22.04.2023 (web resource – CLARIN-PL digital repository).
- Polikarpov, A.A. 2007.** *Komp'yuternyy korpus tekstov russkikh gazet kontsa XX-ogo veka.* [Компьютерный корпус текстов русских газет конца XX-ого века.]. – WEB: https://www.philol.msu.ru/~lex/main_en.htm – WEB: 23.12.2023.
- The National Corpus of the Russian language. 2022.** *Semantičeskaâ razmetka. O leksiko-semantičeskoj informacii v Korpuse.* [Семантическая разметка. О лексико-семантической информации в Корпусе.]. <https://ruscorpورا.ru/page/instruction-semantic/> – WEB: 12.01.2023.

Spis tabel

Tabela 1. Przykład najważniejszych danych w artykule hasłowym w RAS	110
Tabela 2. Dane lingwogeochronologiczne w RAS: RFS i AS	116
Tabela 3. Ścieżka dotarcia do listy AS	140
Tabela 4. Kolumny RAS w wymiarze pionowym (kolumny 1., 2.) i poziomym (3.–9.)	175
Tabela 5. Indeksy a tergo i a fronte z obustronnymi hiperłączami – kolumny 1. oraz 2.	178
Tabela 6. Kolumna 3. w RAS	179
Tabela 7. Kolumna 4. zawierająca kontekst użycia AS	180
Tabela 8. Kolumny 5., 6., 7., 8. przedstawiające metadane: source (title of newspaper), year, RFS and name of the city in this RFS	181

Spis rysunków

Rysunek 1. Eksplozja wzrostu liczby unikalnych słów w dekadzie milenijnej (poszczególne regiony i lata) [oprac. własne]	111
Rysunek 2. Lokalizacje wystąpień słów IAS na mapie regionów [oprac. własne]	138
Rysunek 3. Występowanie AS w prasie lokalnej w różnych RFS w 1996 roku [oprac. własne]	157
Rysunek 4. Występowanie AS w prasie lokalnej w różnych RFS w 1997 roku [oprac. własne]	158
Rysunek 5. Występowanie AS w prasie lokalnej w różnych RFS w 1998 roku [oprac. własne]	158
Rysunek 6. Występowanie AS w prasie lokalnej w różnych RFS w 1999 roku [oprac. własne]	159
Rysunek 7. Występowanie AS w prasie lokalnej w różnych RFS w 2000 roku [oprac. własne]	159
Rysunek 8. Występowanie AS w prasie lokalnej w różnych RFS w 2001 roku [oprac. własne]	160
Rysunek 9. Występowanie AS w prasie lokalnej w różnych RFS w 2002 roku [oprac. własne]	160
Rysunek 10. Występowanie AS w prasie lokalnej w różnych RFS w 2003 roku [oprac. własne]	161
Rysunek 11. Występowanie AS w prasie lokalnej w różnych RFS w 1994 roku [oprac. własne]	161

Rysunek 12. Występowanie AS w prasie lokalnej w różnych RFS w 2005 roku [oprac. własne].....	162
Rysunek 13. Występowanie AS w prasie lokalnej w różnych RFS w 2006 roku [oprac. własne].....	162
Rysunek 14. Mapa miasta Tuła w obwodzie Tuskim [oprac. własne].....	181

